An evaluation of labour market forecasts by type of education and occupation for 1992

AN EVALUATION OF LABOUR MARKET FORECASTS BY TYPE OF EDUCATION AND OCCUPATION FOR 1992

ROA-R-1994/4E

L. Borghans
P. van Eijs
A. de Grip

RESEARCH CENTRE FOR EDUCATION AND THE LABOUR MARKET

University of Limburg
Faculty of Economics and Business Administration

Maastricht, April 1994

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Borghans, L.

An evaluation of labour market forecasts by type of education and occupation for 1992 / L. Borghans, P. van Eijs, A. de Grip ; [transl. from the Dutch]. - Maastricht : Research Centre for Education and the Labour Market, University of Limburg, Faculty of Economics and Business Administration. - ([Report] / Research Centre for Education and the Labour Market, ISSN 0922-8098 ; ROA-R-1994/4E)
Vert. van: Evaluatie arbeidsmarktprognoses naar opleiding en beroep in 1992. - 1994. - (Rapport / Researchcentrum voor Onderwijs en Arbeidsmarkt = ISSN 0922-8098; ROA-R-1994/4). - Met lit. opg.
ISBN 90-5321-125-X
Trefw.: arbeidsmarkt ; Nederland / onderwijs en arbeidsmarkt ; Nederland.

CONTENTS

1. INTRODUCTION

1.1. Background

In 1986 the Research Centre for Education and the Labour Market began to compose medium-term forecasts of the labour market prospects of types of education and occupations, under a contract from the Ministry of Education and Science which was intended to cover the development of an information system of use especially for providing educational and vocational guidance to apprentices and students in secondary and higher education. The first forecasts were used to supplement the labour market module of I See! This was a computerised information system, established by the National Career Guidance Information Centre (LDC), bringing together information from many sources which might be relevant for the choice of a career or course of study. Vocational guidance teachers and others involved in assisting students to make these choices could call up this information via their personal computer and obtain, along with other information on study and vocational choices, an idea of the labour market consequences of the choices which were available. The LDC has now brought out a new information system, 'Traject', which also makes use of labour market information provided by ROA. ROA's forecasts have also been one of the foundations of the LDC's series of brochures for study and vocational guidance. In addition to the Ministry of Education and Science and the LDC, the information system also receives support from the Central Employment Board (CBA), which has used the current data and the forecasts of the information system in formulating policies on employment in general and study and vocational guidance in particular.

As part of this process, the first pilot research project was completed in 1987. This covered current labour market information and forecasts for university education (De Grip, Heijke and Vos, 1987, De Grip *et al.*, 1987a, 1987b and 1988). In 1989 the first forecasts for the full width of the education system were compiled (De Grip *et al.*, 1989 and De Grip, Heijke and Dekker, 1989), covering developments in the period up to and including 1992 for 79 occupational classes and 53 types of education. The forecasts by occupational classes encompassed predictions of the expansion demand and replacement demand, which together comprise the 'job openings'. The forecasts for the various types of education also included predictions of the expected supply, so that a confrontation of demand and supply could be made, on the basis of which a characterization could be given of the expected future labour market situation. The forecasts were supplemented with current data and a number of indicators as regards the occupational classes and types of education which had been differentiated.

The forecasts related to the period 1985-1992, so that the forecast period of these first predictions made by ROA has now elapsed. Bearing in mind the delays in gathering the required data, it is now, for the first time, possible to evaluate these labour market forecasts by occupation and type of education. In fact a first evaluation report appeared in 1991 (De Grip, Heijke and Berendsen, 1991), but since the first forecast period had not then completely run its course, the empirical part of this evaluation was no more than provisional. In addition to this empirical evaluation, however, this report also contained a survey of the methodology, describing the strong and weak points of the approach and possible extensions to the information system.

This second evaluation report will continue along the same lines. It contains three main elements. First of all, the forecasting methodology which was in use in 1989 and the structure of the information system will be described. An overview will be given of the essential assumptions on which these forecasts were based and the important limitations of the system's structure. In fact the forecasting structure which is now being used is in various respects different to the methodology of 1989, so that it is possible that some of the problems noted will already have been by-passed. However this survey of the methodology which was used is still required for a good evaluation.

The second main element is an empirical evaluation of the forecasts which were made, by comparing the prediction with actual changes in the labour market. This empirical testing will be conducted as systematically as possible, but there are a number of points at which the prediction and the actual outcome cannot be correlated, or only to a limited degree. Since it is important, in forming an understanding of the value of a forecast, to have an idea of their (average) accuracy, it must in principle be possible to evaluate them. Therefore, for the components for which no evaluation is possible, an indication will be given of how an evaluation might in the future be carried out, what new data sources would be required, and what methodological changes would be necessary to make an evaluation possible in the future.

The third element of this report will be a discussion of the shortcomings of the approach used at that time, based on the findings of the previous sections. Several further forecasts have been made since the 1989 forecasts for the period to 1992. In 1990 forecasts were made for 1994, again differentiated by types of education and occupation (Dekker *et al.*, 1990, and ROA, 1991). These forecasts were modified in 1991 (Dekker *et al.*, 1992, and ROA, 1992a, 1992b), and in 1993 forecasts were made for 1998 (Dekker *et al.*, 1993, and ROA, 1993a, 1993b). In these new forecasts, as was said earlier, considerable changes have been made to the methodology used in 1989. Therefore, in discussing the shortcomings of the 1989 methods, the modifications which have already been made will be noted, along with the extent to which these may be expected to solve the problems which are described. Finally some conclusions will be drawn as to the significance of these findings for the future, and an overview of the improvements which could be made in the structure of the labour market forecasts will be given.

1.2. Evaluation principles

Although the focus of attention, when compiling an evaluation of forecasts, is mainly on the period in which the forecast was created and the period to which they related, an evaluation is also very important in relation to future forecasting activities.

For the *users of forecasts* of the future labour market, differentiated by education and occupation, it is useful to have some information about the reliability of the forecasts. Borghans (1993) has shown that publicly accessible predictions, as aids in choosing a course of study, have a positive influence on students' choices and therefore on the working of the labour market, provided that students have a reasonable idea of the usefulness of these forecasts. Two things are important in this respect: the first is that the students have a clear concept of the rationale underlying the forecast. The total picture provided by a forecast should be broken down according to the factors from which it is composed, so that the basis on which particular developments are expected is clear. This makes it possible to compare the forecasts with the students' own expectations of future labour market developments and/or various other sources of information. The second requirement is that students have an idea of the average accuracy of the predictions, because this in part determines the degree to which they have to take the forecasts of the information system into account. It is therefore important to check which points the forecasts is reasonably accurate on, and where the uncertainties lie. It is also important to know how the degree of uncertainty is expressed in the way in which the forecasts are published.

A good evaluation of past forecasts is also very important for *those compiling forecasts*. When compiling forecasts a choice must be made between the many possible ways of modelling the labour market. This choice is based on a certain understanding of the functioning of the market. If it was only the quality of the data which determined the quality of the forecasts, the only lesson which could be derived from an evaluation would be a call for more or better data. But an evaluation of the forecasts can also provide new understandings of the applicability of the methods used. This is especially so for forecasts within an information system that is still in the development phase: a fundamental evaluation can also reveal the strong and weak points of the method.

Those who commission forecasts are another group with an interest in the evaluation of the predictions. On one hand this evaluation can provide information as to the priorities which should guide the further development of the information system. On the other hand it is important for those commissioning forecast to know how useful the forecasts are for their target group.

In the light of the central importance of the selection of a model, Granger and Newbold (1986) also argue for a fundamental evaluation as a means of improving the forecasting methodology. They reason that an evaluation should be carried out at two levels: the subjective and the objective. At the subjective level, the forecasts and the actual events are carefully compared and the factors which might have caused the forecasting errors are ascertained, focusing especially on extremely large forecasting errors. It is also possible to determine whether there were incidental causes for the appearance of such exceptional results. One has to ask whether an incidental factor should have been foreseen at the time the forecast was compiled.

One risk of this subjective method, however, is that the evaluation can degenerate into an accumulation of incidents. Where such incidents occur repeatedly, there may be structural shortcomings in the forecasting method which was used. Therefore Granger and Newbold consider the second phase, the objective evaluation, to be even more important. They set out three questions which any objective evaluation should answer (Granger and Newbold, 1986, pp. 277):

(1) are the forecasts better than the available alternatives?

(2) how 'good' are the forecasts?

(3) can the method with which the forecasts were compiled be adapted in such a way that an improvement in forecasting quality would be expected?

These three questions will be the starting-point for the empirical evaluation, but they require some clarifications. The available alternatives to the forecasts must be determined, so as to be able to compare the quality of the forecasts, and the concept of a 'good prediction' must be elucidated. These points will be discussed in chapter 3.

1.3. Goal and structure of the forecasts

Any evaluation of forecasts made within the framework of ROA's information system must rely on an assessment of the extent to which these forecasts accomplish the purposes for which they were compiled. It is therefore important to have a clear picture of the overall forecast structure and of the objectives of the forecasts at the time they were made. This makes it possible to look at both the accuracy of the forecasts and at how far the structure which was employed was suited to the goals which had been set.

Purposes of the forecasts

The 1989 report formulated the primary and secondary goals of the Information System for Education and the Labour Market. "The ROA Information System for Education and the Labour Market has, at least for now, the primary goal of generating information which can be useful in choosing a course of study or an occupation. However this information system could in principle also be useful for capacity planning in the education system, and policies relating to the labour supply, the economic structure and technology, as well as the personnel policy of both the government and businesses." (De Grip, Heijke and Dekker, 1989, p. 1).

Because the second objective is presented here only as an option, and is moreover very general, this evaluation will be related specifically to the first of these objectives. This means that the forecasts and the forecast structure will be examined throughout in terms of their usefulness for those choosing a course of study or an occupation.

The objective of generating information which can be useful in choosing a course of study and for vocational guidance has two implications for this evaluation. First of all, the structure of the information system needs to be suitable, so that the data which are generated can assist in the educational or vocational decisions of an individual student. i.e., the information must:

(1) be relevant for the students;

(2) be relevant at an individual level;

(3) be presented in such a way that students can interpret it properly.

The first requirement, *relevance for students,* means that the forecasts should relate to factors which can be important in making educational or vocational choices. That is, the forecasts must provide understanding of the labour market situation which a student will encounter from the moment of entering the market after the completion of studies. This requirement thus determines the period to which the forecasts must refer — the period in which students will have just left school — and the group in the labour market for which they must be relevant — the school-leavers.

The second requirement relates to *significance at an individual level.* Since the goal is to assist in educational and vocational decision-making, the forecasts must be useable in making an individual choice regarding a particular type of education or occupation. For the student facing the choice of a course of study, it is relevant to know what situation he may face in the labour market at the end of his education. Some developments may be very relevant, in a general sense, for a description of the labour market, but only useable for an individual if they are translated to the individual level. For instance, the total growth in employment for a particular educational category is interesting for those making policy decisions, but for a student it is more important to know what his or her individual chance of getting a particular sort of work at the end of the course may be.

The third requirement for the structure of the forecast system, if it is to be useable for educational and vocational guidance, is that *the information must be presented in an interpretable form.* If they are to be useable for students, the forecasts must for expressed in terms which are comprehensible for someone who is not entirely adept in labour market interpretations.[1] The forecast results should, as far as possible, be expressed in generally understood concepts with a minimum of statistical or economic jargon. A translation is especially important for statistical judgments regarding the reliability of the forecasts. Improvements in the ease of interpretation of the labour market data can however come in two ways: it may also be sensible to improve students' understanding of the functioning of the labour market (as is at present being considered as part of the reform of the basic secondary education curriculum) so as to improve the comprehensibility of labour market information.

A second consequence of the objective of generating information which is useful for study and occupational choices is that the empirical evaluation will also be made from this standpoint. This means that the evaluation criterion which is used in determining the magnitude of forecasting errors must in the first place show what consequences these forecasting errors have had for individual occupational and vocational choices. The selection of a criterion for the evaluation is discussed in section 3.3.

Structure of the forecasts

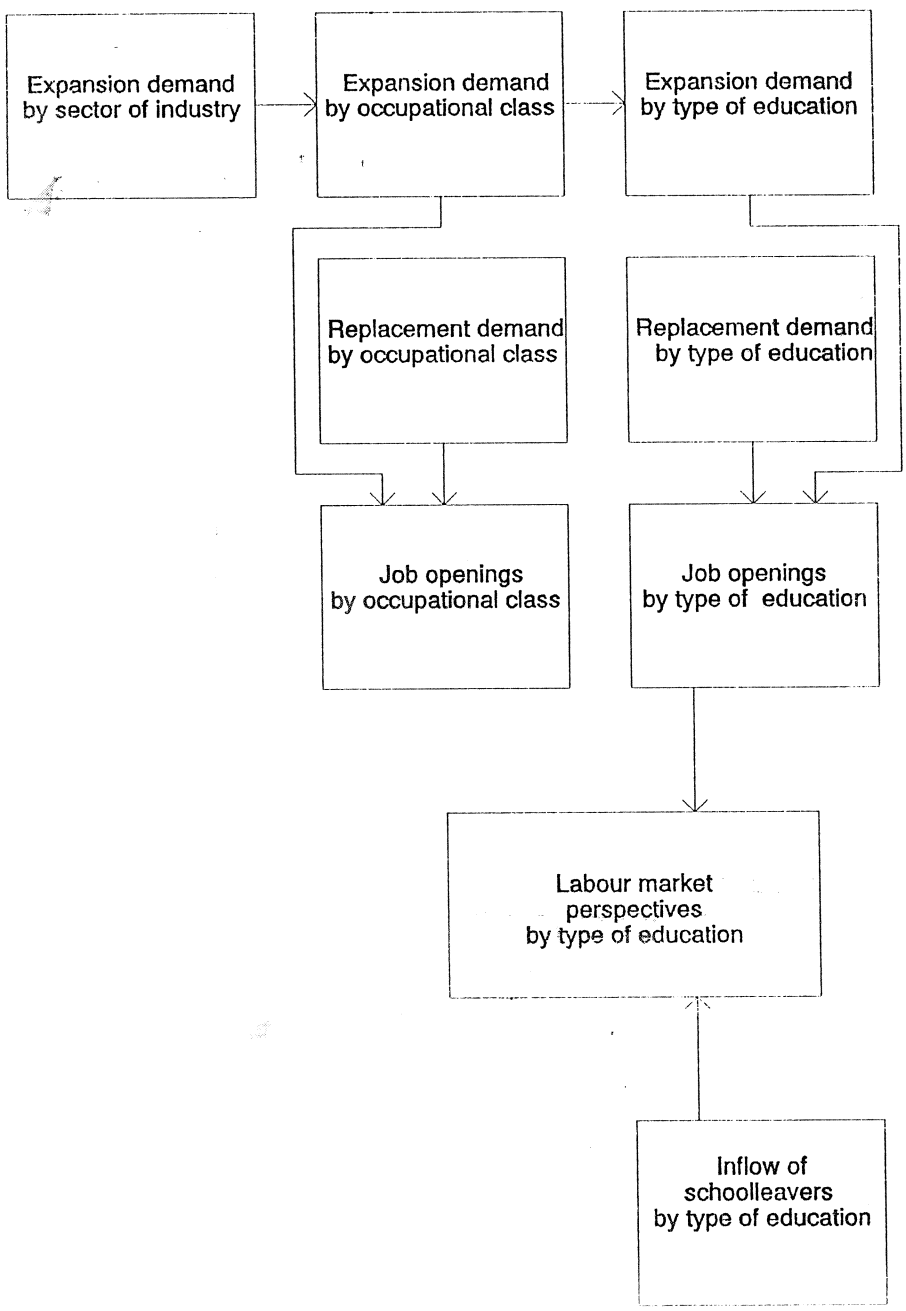
Figure 1.1 gives a schematic overview of the structure of the forecasting method. On the demand side, the forecasts of employment in economic sectors which are obtained from the Central Planning Bureau (CPB) provide an external source of information. At that time, these forecasts covered 22 economic sectors. The predicted numbers employed in these economic sectors are then

---

1. The data from the information system is in fact not provided directly to students, but the I See! system is made available to vocational guidance officers in schools. One reason for this has been to foster correct interpretations of the predictions.

translated into the employment in 79 occupational classes. This predicted level of demand for occupations, when compared with the situation in 1985, yields the expected expansion demand for each occupational class. This is supplemented with a forecast of the expected replacement demand. Together, the expansion and replacement demand comprise the expected number of job openings. It is assumed that it is the number of job openings which is the relevant quantity for educational and vocational decisions. New entrants to the labour market cannot in practice simply crowd the people already working out of their jobs. On the basis of a breakdown of past flows into the labour market, a calculation is also made of how many of these job openings are available for school-leavers.

Figure 1.1. Structure of the forecasting model of the Information System on Education and the Labour Market



In addition to these forecasts of the number of job openings per occupational class, the information system also contains current data and indicators for each occupational class. The current data covers the number of workers in a particular occupational class, a breakdown into the component occupational groups, the age distribution of the workers, a breakdown by educational categories and types of education, economic sectors and sub-sectors, and the proportion of self-employed workers, plus the trends in these figures. The underlying idea is that such information can contribute to widening the horizons for those receiving educational and vocational guidance and in various ways give some understanding of a number of relevant characteristics of the labour market for a particular occupational class.

The indicators relate to the assimilation rate of workers who are under 30 years old, the dispersion of the occupations among the various economic sectors and the sensitivity of employment to the state of the business cycle. The assimilation rate shows how many young people work in an occupation, as compared with other occupations. The dispersion over economic sectors and sensitivity to cyclical forces are both 'risk indicators'. The dispersion indicator shows how widely the employment for workers in this occupational class is spread over various economic sectors. If there is a wide dispersion, any unexpected change in a particular economic sector will have relatively little influence on the developments in that occupation. Wide dispersion would also indicate that, if demand in one of the economic sub-sectors should decline, there would probably be possibilities of work in other economic sectors. The sensitivity to cyclical forces shows the extent to which employment for the occupation rises and falls with the fluctuations in the business cycle. High cyclical sensitivity means, in the first place, an increased chance that the labour market prospects may deteriorate at some time later in a worker's career. It also indicates that there is a greater likelihood that the forecasts will not be realized.

The expansion demand for each occupation is translated, by means of a distribution model, into the expansion demand per type of education. The replacement demand for the various types of education is calculated separately, because this cannot be derived from the replacement demand per occupation. Together, the replacement and expansion demand make up the total demand for new entrants with a particular educational background.

In addition to the forecasts of demand, a forecast is made for each type of education of the expected supply of school-leavers entering the labour market between 1985 and 1992. These forecasts are based on the Skill, Rhobos and Worsa forecasts,[2] supplemented with additional data on part-time education and courses outside the regular, government-supervised education system (i.e. 'non-regular' education). The total supply of new entrants, divided by the total demand for new entrants, is called the 'labour market indicator' (LMI), which has since been replaced by the 'indicator of the future labour market' (IFL). This ratio gives an indication of the tensions between demand and supply in the period under consideration. Because the model which has been used

2. Skill, Rhobos and Worsa were estimates used at that time by the Ministry of Education and Science for the flows entering, within, and leaving the education system. The Skill forecasts related to all training levels, Worsa covered a break-down of the Skill forecast covering University Education, and Rhobos was a break-down of the Skill forecast for Higher Vocational Education alone.

takes no account of possible adjustments in the labour market in response to these tensions between demand and supply, this indicator of tension must not simply be regarded as the expected over-supply or shortage. Naturally both the demand and supply sides of the labour market will to some extent adjust in response to these tensions, so that, for example, an over-supply of new entrants with a particular education will not necessarily be expressed in high unemployment, but could also lead to stagnating wages, an increased probability of having to find work outside of the intended field of employment, or other deteriorations in the labour market situation.

In addition to the forecast of the future labour market situation, supplementary data is provided for the types of education, just as for the occupations. In the first place a picture is again given of the current labour market situation. This information relates to the number of workers, the most important occupational classes and occupational groups and the unemployment (as well as the proportion of unemployment which relates to school-leavers), with an indication of the trends in these figures. Furthermore, an indication is given of the dispersion of employment across occupations and economic sectors. Like the indicators for occupations, these dispersion indicators give an impression of the robustness of the forecasts, but also show in a more general way the extent to which the choice of a particular type of education will leave students dependent on the labour market situation for a particular occupational class or a particular economic sector.

For educational and vocational guidance purposes, the quantitative data of the forecasts, current data and indicators are all transformed into qualitative characterizations. On the basis of the values which have been established for the variables, a classification is made in each case on a five-point scale, on which the ranges are characterized as 'very low', 'low', 'average', 'high' or 'very high'. The intention of these qualitative characterizations is to render the quantitative figures of the forecasts, indicators and current data more accessible for people who are not accustomed to dealing with such figures. In the first place they do not have to understand the measurement units in which the variables are measured. In the second place, this method gives an immediate relative characterization, so that it is not necessary to examine the dispersion of the variable. In the third place, the division into five intervals produces a characterization which gives a less exact impression than the figures themselves, avoiding the suggestion of accuracy to the last decimal point, and giving a certain bandwidth to the labour market forecasts in particular.

1.4. Structure of the report

The report proceeds as follows: chapter 2 discusses the structure of the information system, examining how the forecasts were drawn up and how these predictions, supplemented by the current data and indicators in the information system, provide a picture of the labour market developments for the various kinds of education and occupation, and especially the future labour market prospects of the types of education. This begins with the overall structure of the forecasts and the data which has been used, followed by a discussion of the various components of the information system and, finally, the qualitative characterizations and risk indicators.

In chapter 4 the components are, so far as possible, empirically evaluated, based partly on the findings in the previous chapter regarding the structure of the information system. An evaluation methodology for the purpose is described in chapter 3, which serves here as a starting point. On the basis of this empirical evaluation, the overall predictive quality of the information system is measured, and we examine where the methods used were particularly inadequate. This will enable us to examine which methodological changes would be likely to improve the quality of the system.

In chapter 5, finally, the findings of both the subjective and empirical components of this report will be brought together on the basis of three tables which provide an overall view of the evaluation. The first table contains some key figures which indicate the predictive quality of the components of the forecast model. The second table contains an overview of aspects of the structure of the information system which, according to the findings in this report, should be adapted, or have already been adapted. The third table contains a similar list of possible modifications as regards the forecasting method.

2. THE STRUCTURE OF THE INFORMATION SYSTEM

2.1. Introduction

Because the primary objective of the forecasts is to delineate the future labour market position of school-leavers, we will select some particular aspects of the future labour market which are relevant in this regard. This chapter will survey the most important principles of the structure of the information system and the methodology in which these principles have been expressed. The goal of this survey is to give an overview of the strong and weak points of the information system, so as to clarify which components deserve more attention during the further development of the forecasting method. Moreover, this analysis of the structure and methodology will highlight a number of points to be considered in the empirical evaluation.

2.2. Overall structure of the forecasts

Labour market prospects versus chance of finding work

The hub of the forecasts in the information system is the confrontation of demand and supply in the labour market. This confrontation — at the level of types of education — brings together the various components of the information system to produce the LMI ratio (now called the IFL ratio, the 'indicator of the future labour market'). Despite its key position in the information system, it is difficult to give an unambiguous interpretation of the results of this confrontation of demand and supply.

Unemployment is the worst and certainly the most visible effect of a bad match between the education system and the labour market. It is therefore not surprising that any prediction of bad labour market prospects is generally interpreted as a prediction of a low probability of finding work with the type of education concerned. The assumption is often that, if the supply exceeds the expected demand, the difference must be expressed in a corresponding increase in unemployment. On the other hand, a demand which exceeds the supply is thought to lead to a corresponding number of unfilled vacancies.

Nevertheless, the primary position given to unemployment when considering mismatches between the education system and the labour market is erroneous. Unemployment is certainly not the only way in which an expected discrepancy between demand and supply may be expressed. If there is some tension between the quantity of labour on offer and the demand for it, the labour market may adapt to the tension in various ways. According to the neoclassical approach, such a tension would simply lead to wage adjustments. If the supply exceeds the demand for a particular education, in this theory, the wages for people with that kind of education should fall. Although this would have no effect on the supply — because the school-leavers who are at that moment entering the labour market have completed their education — wage reductions can nevertheless increase demand for this labour, which is now cheaper, so leading to an accommodation of the demand to the supply.

In reality, however, many frictions between demand and supply are not resolved by wage adjustments. If supply exceeds demand, this puts those offering labour under pressure. In such a situation, the increased competition on the supply side will drive them to accept less favourable working conditions and will encourage them to switch to jobs which are less attractive for them (for example, jobs at a lower level). Finally, an excess of supply naturally also means an increased chance of not being able to find any work or of having to look for work for longer than would normally be the case. In general excess supply will thus mean that those offering labour in the market are in a weaker position than those on the demand side. A supply which exceeds the demand therefore means, in one respect or another, bad labour market prospects. Even in the theoretical case in which all adjustments occur via the wage mechanism, employees will be worse of where supply increases, because they will have to accept lower wages.

The fact that a discrepancy between demand and supply need not, *per se,* lead to unemployment or unfilled vacancies can be clearly seen from the 1989 forecasts of demand and supply by types of education which are being evaluated in this report. If all the excesses and shortages of labour with a particular educational background are brought together, this would, granted this premise, have resulted in unemployment of 736,000 people and 299,000 unfilled vacancies, considering only the effects of the flow of school-leavers entering the labour market. This would be 42% and 17% of the total flows, respectively.

Adjustment processes

Discrepancies between demand and supply will therefore lead in some way to adjustment processes in the labour market. Borghans and Heijke (1993) show that these adjustments are in principle unfavourable for whichever side of the market has an excess. Therefore the difference between the predicted demand and the predicted supply, where the adjustment mechanisms are not taken into account in the forecasts, gives an indication of the extent of the adjustment effects combined with the unemployment or vacancies which can ultimately be expected. The difference between demand and supply is thus certainly an indication of the labour market prospects. Van Eijs (1994) notes in this respect that the extent of adjustments can vary, for a given magnitude of discrepancy between demand and supply, for each type of education. Depending on the elasticities in the adjustment mechanism, a small movement will suffice to bring the demand back in balance with the supply for one type of education, while a large movement will be required for another type of education.

On the basis of this reasoning, we can conclude that a general indicator of tension, if it is to serve as a measure of the labour market prospects of school-leavers, must in principle be based on forecasts of demand and supply without considering any adjustments. It would also be very helpful if the prediction covered not only the tensions between demand and supply, but also enabled one to form an opinion of which adjustments would be expected in a particular case.

As regards the first point there is a problem, at least for the forecasts being evaluated here, in that the adjustment processes are in part included in the forecast model. In the model for the types of education, the demand for particular types of education is in part determined by the composition of the supply by educational levels. The explanation of this adjustment mechanism is that, if the supply of those with higher education increases, employers will revise their demand. According to the line of argument which was described above, however, this adjustment may already mean a worsening of the position of those with higher education. Because the adjustment has already been included in the model, it has no further effect on labour market prospects which are expected on the basis of the confrontation between demand and supply. For this reason, in the recently compiled forecasts up to 1998, the influence of supply factors is not considered when determining the demand for types of education (Borghans and Heijke, 1993). Moreover, this new educational model explicitly recognizes the fact that adjustment processes for a particular type of education can influence the labour market prospects of other types of education, via substitution processes and the tendency of one type of education to crowd another out.

This explicit modelling of the adjustment mechanisms has the advantage of making it possible to predict how a discrepancy between demand and supply will make itself felt. However, because the prediction of general labour market prospects would be expected to be much more achievable than a precise forecast of the concrete way in which this will be manifest, it would appear sensible to at least forecast these general prospects, with a prediction as to how it may be expressed added as a supplementary detail. The informative value of the information system could therefore be increased by giving an indication of the expected unemployment, degree of under-utilization, wages, etc. for a particular type of education. This could be based on information from surveys of school-leavers, using the forecasts of the expected labour market prospects of that type of education.

Level versus change

The fact that forecasts for labour supply and demand are always based on changes in demand and supply in the past implies that, if the adjustment processes are not incorporated in the model, the forecasts are based on the assumption that the circumstances under which workers with a particular educational background are hired will remain the same as those of the recent past. The forecast of the labour market prospects must then also be regarded as a forecast of the change in these circumstances. The net effect of this may be that a type of education offering very favourable conditions in the current situation, but for which bad prospects are forecast, may still offer school-leavers better conditions when they enter the labour market than a type of education with favourable prospects, but very bad conditions at the moment. The prospects are, as it were, relative to the conditions which a student could expect for that type of education. In concrete terms this means, to give an extreme example, that lower technical education is not immediately attractive for a potential university-level student of engineering, just because the prospects for Lower Technical Education are good and those for Engineering are only average.

For this reason the inclusion of data on the current labour market position of a type of education is very important for the usefulness of the forecasts. The expected prospects must in fact be related in each case to the current situation. Unemployment may be an exception to this rule that the labour market prospects indicate the expected change in the situation. Because unemployment can be seen as the effect of delayed adjustments to an excess of supply, it would be expected that, given average labour market prospects in the longer term, the unemployment rate would ultimately stabilise around the average unemployment rate. The unemployment rate itself, and not the change in the rate, would then be related to discrepancies between demand and supply.

The level of the confrontation

In the information system, the supply of labour is primarily dealt with in terms of types of education, while the demand for labour is in the first instance considered at the level of occupational classes. But if a confrontation of demand and supply is to be made, one level must be chosen for both measurements. In the structure of the Information System on Education and the Labour Market, this is the level of types of education. The translation of demand per occupational class into demand per type of education is, however, not without its problems. While supply is clearly related to a particular type of education — in fact it takes a long time to retrain in another type of education — the demand for labour is certainly not always limited to a particular type of education. Often many types of education might be considered for a particular occupation. One may well be more suitable than the other, but there is no reason to make this an absolute limitation. The reverse of this is that people with a particular type of education can find places in various occupations, but, most importantly, it also means that the jobs which are in principle open for this type of education are also available for people with other educational backgrounds. Borghans and Heijke (1993) introduce a model in which the competitive interaction between the types of education is explicitly included. In Borghans (1992) and Van der Velden and Borghans (1993), the competitive position of the types of education is covered in more detail.[3]

It is however also possible to make a confrontation of demand and supply at the level of occupational classes. The supply by types of education would then have to be translated into supply per occupational class. In this case the discrepancy between demand and supply would indicate, in terms similar to the discrepancy for types of education, what changes might be expected in the recruitment problems for particular occupations. The reason that a confrontation at the level of types of education has been chosen in the information system is that this is the most relevant dimension for people choosing a course of study. It would in fact be meaningful to compare demand and supply for occupational classes, in addition to the confrontations for types of education, so as also to look at the forecasts from the information system from the demand side of the labour market.

Competition for jobs

It has been said above that it is important, in determining the labour market prospects of types of education, to know which types of education compete with each other for the same jobs.

3. Comparable data has also been published in the report on the incorporation of data from school-leaver surveys in the information system (Wieling, De Grip and Van der Velden, 1993).

Furthermore, to focus the forecast of labour market prospects on school-leavers, it is necessary to identify the jobs for which school-leavers are eligible, and the groups who compete with the school-leavers for these jobs. By clearly defining a specific segment of demand and supply, the ratio between the two can be determined.

In making the forecasts for 1992, it was assumed that school-leavers are eligible for the new jobs (expansion demand) and the vacancies created by the departure of workers with the same education or the same occupation (replacement demand). It is also assumed that they face no competition in securing these jobs from other job-seekers. The implicit supposition is that students cannot take jobs from people who are already working. If those already working try to get a new job, new entrants do not in principle suffer, because the net effect of this mobility cannot reduce the number of job openings. However if a working person is forced to find another job, he or she has, according to the assumption, a better chance than the new entrants, who are left with only the net increase in jobs. There is no competition on the supply side with other groups: in other words, the unemployed and older workers re-entering the market apparently are not in competition with school-leavers.

Since the forecasts of 1990, competition from the unemployed is also taken into account on the supply side. It is assumed that only those who have been unemployed for less than a year are a serious threat to the position of school-leavers.

2.3. The data

Three sources of data were very important in compiling the forecasts for the information system, and for the associated current data and indicators. These were the Employee Census (AKT) of the Central Bureau of Statistics (CBS), the estimates of student numbers which are prepared by the Ministry of Education and Science and others, plus data from the Central Employment Board as regards the numbers of registered unemployed. These three sources of data will be discussed separately in this section. Forecasts of employment by economic sector and a number of other exogenous variables derived from the CPB were also used, but the CPB data will be discussed along with the empirical evaluation (in chapter 4).

The Employee Census and Workforce Census

The principle source of data for the demand module of the information system has been the Employee Census of the CBS, which has now been replaced by the Workforce Survey (EBB). The Employee Census provides data on the numbers of people working in each economic sector and occupational class, and from each type of education. This data is used to forecast the spread of the demand per economic sub-sector over occupational classes and types of education. The continuity of the data is important for the forecast and its evaluation.

When compiling the forecasts, the Employee Censuses of 1979, 1981, 1983 and 1985 were available. Versions of the Employee Census from before 1979 were not useable because they were based on a different educational classification. The forecasts had therefore to be based on only four historical observations. Annual EBB data is now available from 1988 onwards for economic sectors and occupational classes and from 1990 for types of education. However the structure of the EBB survey differs from that of the Employee Census, so that the two sets of data could not be combined without some difficulties.[4] The principal change was in the definition of working people, a change which has especially large consequences for the small part-time jobs. There has also been an adjustment to the way in which the departments within a business are classified under economic sectors.

For the forecasts utilizing both the Employee Census and EBB data, the main change concerns the minimum number of working hours in the definition of the workforce. In the Employee Census, the question asking whether people have paid work did not contain an explicit lower limit for the number of hours worked. In the EBB there is an explicit link with working hours. On the basis of this data, the CBS now counts everyone who has at least 12 hours per week of paid work as working. However it is also possible to derive the number of those working at least 1 hour per week from the data. In principle this latter definition should match the definition which is implicit in the Employee Census, but because the question is asked in a different way, a large number of people with small part-time jobs are not registered as working by the Employee Census (Bierings, Imbens and Van Bochove, 1990). As a result the Employee Census and EBB data cannot be directly converted to one another.

Despite the changes in structure, the EBB can be correlated with the Employee Census reasonably well, so that historical observations were available for the forecasts which have recently been compiled for 1998 for the class 8 occupations and the group 6 types of education. Because the CBS will soon switch to a new occupational classification, it is not yet clear whether the Employee Censuses will remain available for the research. It will very likely be impossible to recode the Employee Censuses on the basis of the new occupational classification. As a result, the available time series will again be considerably shortened.

In fact there have been a number of classification changes since the publication of the first forecasts in 1989. The classifications of economic sectors, occupational classes, and educational types to which the forecasts related are based on the customary classifications used by the suppliers of the data. For the economic sectors this meant the classification of economic sectors used by the CPB for its medium-term estimates. In 1989 a classification into 22 economic sectors was used. Since 1990 ROA has used a classification into 14 economic sectors, stemming from the Athena model of the CPB. A number of small changes have since been made to this classification system.

For the occupational classification system, ROA has decided to discard the very obsolescent system of CBS occupational classes which was used at that time. It was concluded that the CBS classification was not appropriate for use in a description of the labour market process at this low

---

4. Problems regarding this linkage are discussed in Dekker and De Grip (1993).

level of aggregation. The CBS occupational classes were too heterogeneous as regards the educational background of the workers, and were not sufficiently recognizable for the users of ROA's information. Therefore a new classification system has been constructed, in which CBS occupational groups are clustered on the basis of the educational profile of the workers in the occupational group.[5]

The classification of types of education which was used in the forecasts which are to be evaluated here is an aggregation of educational categories from the standard education classification of the CBS at the 3-digit level. This classification has remained in use to the present day, but has undergone small changes in particular points in response to the problems encountered with the classification in practice. These changes in the classification system have however always been designed in such a way that it remained possible to continue to use the Employee Census and the EBB results from all the available years. This ability to transpose the figures will remain important in any new changes which may be made.

A second aspect which is very important in relation to the data used is the quality of the sources. Because both the Employee Census and the EBB are random samples of the Dutch workforce, both contain a sampling error. Mistakes in the sampling process, and in the processing of the questionnaires, will introduce additional errors which reduce the quality of the data. Because the EBB is based on a sample of 1.1%, while the Employee Census was generally based on 3% of the population, the sampling error has increased considerably. Moreover the CBS has for some time now only provided figures rounded to the nearest thousand, for published data, and to the nearest 500 for data sets for research purposes. This increased the margin of error in the data.

However it must be said that the EBB, in contrast to the Employee Census, is published annually, so that the quantity of information available per year has been only slightly reduced. Certainly, so far as the reliability of time series analyses goes, the annual EBB data is preferable. Especially if one is using the data in a statistical analysis (time series or cross sectional analyses), the errors in the data are acceptable. However these measurement errors must be taken explicitly into account in designing the model, to a greater extent than is now done, to ensure that the measurement errors do not dominate the results. Furthermore, the convention of rounding figures provided for research purposes off to the nearest 500 introduces an unnecessary reduction in quality where the data is used as input in research. It would be preferable if data for use in analyses was available without any rounding-off, although the published figures could in that case certainly be rounded-off.

In addition to the rule of the CBS that the available data has to be rounded off, there is also a rule that no figures which relate to less than 5,000 people may be made available. For internal research purposes this limit is 2,500. Where three-year averages are used, the lower limit for published figures is lowered to 2,000. Although the used of three-year averages for research purposes leads to loss of information, this may be a suitable alternative for small occupations and types of

---

5. De Grip, Groot and Heijke (1987) contains the initial analyses which provide the basis for this new classification system. The new system was definitively established in Dekker, De Grip and Van de Loo (1990). In Dekker and De Grip (1992) the CBS and ROA systems of occupational classification are compared.

education, rather than using only the current EBB data.

Flow data

The data as regards the flows of school-leavers, which is used as a forecast of the new supply entering the labour market, is in the first place based on the estimates of student numbers prepared by the Ministry of Education and Science. In 1989 these estimates were 'Worsa', for University Education, 'Rhobos', for Higher Vocational Education, and 'Skill', for the whole regular education system. Since then, the three forecast models have been replaced by one set of estimates, known as 'Reference Forecasts' covering the whole regular education system. These estimates are based on data from the comprehensive student count carried out by the CBS. Because the forecasts of supply for the information system are directly based on the 'Reference Forecasts', the CBS's comprehensive student count is not a direct source of data for the information system.

In addition to the estimates for regular full-time education, numerous sources of information have been used for non-regular components of the education system. This includes data on recognized forms of correspondence education, company training programmes and the training programmes run by the Employment Board. Separate sources were also consulted for part-time education. The information from these sources is often not readily accessible, and may be unreliable or classified in a way which does not match the CBS classification system. The use of multiple sources of data also entails the risk of double-counting. It would therefore be an improvement if student numbers could be collected and published for non-regular education as is done for regular education.

Unemployment data

The current data in the information system includes figures for the unemployment for the various types of education. This data is based on the Employment Board's unemployment register. The number of unemployed was counted on the basis of the situation in mid-April 1987. The decentralisation of the Employment Board since then means that no figures on unemployment are now available from this source.

In place of these figures for registered unemployment, the information system now uses the unemployment statistics from the school-leaver surveys, the Higher Vocational Education monitor and RUBS. However, these sources only provide information on the unemployment among school-leavers, and even then not for those with University Education or primary education alone. A good source of unemployment statistics over the full width of the labour market is however very important, both to be able to dispense good current information on unemployment and also to enable unemployment to be explicitly modelled in the further development of the forecasting methodology.

2.4. Expansion demand per occupation

As was already noted in section 2.2, regarding the overall structure of the forecasts, the CPB forecasts of employment per economic sector are the input for the forecasting model for the expansion demand. In the first instance, these employment forecasts by economic sectors are converted into employment forecasts by occupational classes. This conversion includes the transformation of employment measured in labour years into employment in terms of numbers of individuals, on the basis of P/L (people/labour years) ratios for each economic sector. These ratios are predicted by the CPB.

The forecasts of employment levels for each occupational class (Dekker, De Grip and Heijke, 1988) are entirely determined by demand factors specific to economic sectors, with developments on the supply side of the labour market playing no role at all. In the CPB forecasts, however, the demand for labour at the macro level is partly determined by supply factors. In the first place, the supply has a direct influence on the total demand and, in the second place, the supply influences wage levels which in their turn have an effect on demand.[6] The structure of the model means that demand at the level of occupational sub-sectors, disregarding the influences at macro level, must be interpreted as the *ex ante* demand, i.e., as the demand which would have existed if the balance between the labour demand and supply had remained the same. Another way of interpreting this approach is to suppose that demand at the level of occupational sub-sectors is just not influenced by supply factors, i.e., that the demand per occupational class is fixed for the employers and they are able only to decide, as a result of disequilibriums in the labour market, to change the types of workers who engage in these occupations. In other words, the organisational structure does not adapt to scarcity problems in the labour market, it is only the educational background of the people who perform the required activities which can vary. The interaction between demand and supply then occurs only in the educational model.

Although the published reports concerning the structure of the information system in 1989 were not explicit on this point, it will be presumed in this evaluation that there is an implicit assumption that the demand per occupational class is independent of supply factors and that at this level the demand *ex post* — i.e., after adjustments in response to disequilibria have been made — is thus equal to the demand *ex ante.* This assumption is more plausible because this distinction is explicitly made as regards the analysis for types of education, and it is explicitly supposed that the *ex post* demand can differ from the forecasts because of adjustment processes in the labour market.

The distinction between *ex ante* and *ex post* demand is crucial for the evaluation, because the *ex ante* demand cannot simply be compared to the demand which actually eventuated. If the demand forecast is interpreted as *ex post* demand a direct comparison can be made, but this entails two important suppositions. The first is that employers do not respond to labour scarcity by shifting their demand between the various occupational activities, so that the demand per occupational class stays the same. The second is that, although employers have a preference for particular types of education when filling vacancies, they will employ the same number of people if scarcity problems force them to switch to other types of education (i.e., they will not increase the quantity

---

6. See, for example, the figure in CPB (1990, p. 4) for the Athena model or the figure in Eijgenraam and Verkade (1988, p. 4) for the Beta model which was used at that time.

of people to compensate for them being less suitable).

Structure of the occupational model

In transposing the employment forecasts per economic sector into employment forecasts per occupational class, two components play an essential role: changes in the employment levels per sector and changes in the occupational structure of employment within each sector. One can begin with the occupational structure of each economic sector in the past, and so forecast the employment per occupation by simply calculating the results of changes in employment levels per economic sector. This approach is called the fixed coefficients model. Models based on simple fixed coefficients are widely used, for example by the Bureau of Labor Statistics (BLS) in the United States.[7] This method is however not able to predict changes resulting from technological development, because it is assumed that the existing occupational structure of each sector remains unchanged. Bishop and Carter (1991) show that the BLS has been unable to adequately forecast the growth in particular occupations, partly as a result of this limitation.

If better predictions are to be made than is possible on the basis of the fixed coefficient method alone, an explanatory model of the occupational structure is required as a second component. Such a model was utilized for the forecasts of the Information System on Education and the Labour Market in 1989. The shares of occupational classes in the total employment of an economic sector are made dependent on explanatory variables. In the forecast model used in 1989, the explanatory variables are the trend, the ratio between investments and added value, the utilization rate and the share of computer specialists (CBS occupational class 08) in the employment in the sector. The model also contains a different constant term for each occupation in each economic sector, known as the economic sector constant.

In using such a model in practice to explain the shares of occupational classes in the total employment of economic sectors, three problems are encountered. The first relates to the choice of the explanatory variables. The second is the specification of the equation which is to be estimated and which then provides the basis for the forecast of the occupational structure. The third problem area arises in estimating the equations, due to the limited quantity of data available. These three problems will be considered in sequence.

Explanatory variables

In the 1989 occupational model, and in later studies in this area (Peeters, 1990 and Borghans and Heijke, 1994), it was evident that, of all the explanatory variables which are used to predict changes in the occupational structure of types of education, the trend contributes the largest part of the explanatory power of the model. The more economic explanatory variables listed above can

7. See BLS (1988). Van Eijs (1993) provides an overview of the historical development of manpower forecast methods. Van Eijs (1994) evaluates the forecasting methodology of a number of important Western institutes.

explain only a small part of the shifts in the occupational structure. This brings with it the disadvantage that the model as used gives only a limited picture of the causes of changes in the occupational structure. A more important disadvantage, however, is that the exclusive reliance on the trend in the forecasting method always entails a risk that this trend may be automatically, but incorrectly, incorporated in the forecast because it is unclear what the cause of the trend was.

Finding good explanatory variables therefore remains very important. This can be achieved, in the first place, by further examining which factors are likely to make a significant contribution to explaining the changes in the occupational structure. Because many of the factors which might be important cannot be measured, or can only be included very indirectly on the basis of the available information, it may also be necessary to seek out additional databases. In the current structure, all the explanatory variables in the model operate at the level of economic sectors. The occupational model does not yet include any explanatory variables at the level of occupational classes.

In the meantime the trend, as has been noted, remains an important explanatory variable in the occupational model. The extrapolation of trends will always play a role, because it is not realistic to expect that all the processes which go on in the shifts between occupations could be declared on the basis of an economic model.

The 1989 forecasts were based on a time series which was built up from only four Employee Census surveys covering a period of 7 years. A time series of 13 years, with observations at 9 different times, is now available for the explanation of the occupational structure. The increased length of the time series means that large consistent shifts covering the whole observation period are less frequent. This expansion of the time series naturally offers opportunities for better forecasts, but, on the other hand, also requires that adjustments be made in the method of dealing with any temporary trends. What has to be avoided is that trends which have in fact already levelled off are extrapolated regardless, or that trends which only appear in recent observations are not fully recognized in the longer time series. This means that recent observations have to be treated in a different way to older observations.

The specification of the model

A second point that must be filled in in detail in an occupational model is the specification of the mutual relationships between the share of an occupation in the employment in an economic sector and the explanatory variables. The specification which is chosen largely determines how the developments which are observed are extended into the future (see Borghans and Heijke, 1994).

The choice of the model specification also has considerable repercussions on the treatment of the most recent observation in the available time series. In the 1989 model, the occupational shares were regressed against a number of explanatory variables. Naturally, the actual shares differ from the regression line which was found. If the regression line is nevertheless used for the forecasts, the information which may lie concealed in the difference between the most recent observation and the occupational share which was estimated for that year is ignored. The divergence in the most recent observation may be due to accidental factors, but may also be due to the effects of factors which have not been incorporated in the model. It is very likely that any such factors would also play a role in the immediate future and may thus be important for the forecasts. This gives rise, in econometric terms, to autocorrelation of the error term. But in the model which was at that time employed, no account was taken of this mechanism. Since then, however, the model has been changed to incorporate this aspect. In Peeters (1990), a lagged endogenous variable was included as an extra explanatory variable, while Borghans and Heijke (1994) chose to specify a growth model. Both of these specifications of the model take the recent value of the occupational share as the starting-point for the forecast, so that the information value of this recent observation is taken into account. This also means that it is no longer necessary to incorporate separate constants for each economic sector in the model. This is also an advantage in the light of the growth of the available time series. A similar approach is advocated by Clements and Hendry (1993) and Wilson (1994).

Estimating the model

A third important point in filling in an explanatory occupational model is the way in which the model is estimated. Since the available time series is fairly short, it is not possible to estimate a separate explanatory model for every occupation. Therefore, in the occupational model being evaluated here, all those working with a given occupation, in whatever economic sector, are considered as one. This implies that an estimate is made for an occupation, using the same equation to cover the development of that occupation in every economic sector. However a distinction is made between the agriculture and industrial sectors on the one hand, and, on the other hand, the profit and non-profit services sectors, because of the very divergent character of these two sector clusters. However even with data pooling, a number of parameters of the model could not be accurately estimated. This can be seen from the sometimes extreme estimates for the parameters. It would appear that the estimation errors dominated the parameter estimates because of the lack of adequate data. The use of the estimates would have led to instable predictions. Therefore, in the occupational model that was then in use, it was decided to apply a model selection, between three variant models. The first variant is a model with explanatory variables and a trend. The second variant is a model with explanatory variables and no trend, and the third variant is a model without any explanatory dimension, i.e., the fixed coefficients model. For each occupation, in each of the two divisions of the labour market — agriculture & industry on the one hand and commercial & non-commercial services on the other hand — one of these three types of model is chosen. For occupations which are represented in only one economic sector within one of these divisions of the market, the variant without explanatory dimension was chosen at once. For the other occupations, the first consideration is whether the observed trend is significant and whether its sign is plausible. If not, the trend is omitted. Then the explanatory part of the model is examined *in toto.* If this also is not significant, resort must be had to a model without explanatory variables.

One consequence of this selection process is that — for both agriculture & industry and commercial & non-commercial services — a trend term is incorporated in the forecasts in only ten cases. This appears to contradict Bishop and Carter (1991) who emphasise the great importance of trends for good occupational forecasts. However there is no assurance that it would have been desirable to incorporate the trend in every case. Because the time series which were available were very short, the quality of the trend-based forecasts would probably have been low, with a good chance that these estimations would have led to bad predictions. In any case the dilemma as to whether or not to incorporate a trend variable has been solved because the available time series has grown longer and also because a random coefficients model has been adopted, which offers a middle way between using and not using a trend. This new approach uses a weighted combination of the forecasts based on the variants with and without a trend, with the weighting depending on the reliability of the estimations. This approach avoids the all or nothing strategy, under which doubts about the reliability of the trend meant that no trend could be incorporated at all. This approach also means that the parameters of various occupations and various economic sectors do not have to be the same. This allows more room for variations between occupational classes and economic sectors.

Conclusions

In brief, the occupational model used in 1989 has, apart from the fundamental suppositions concerning the nature of demand at occupational level, two possible shortcomings. First of all, little use was made of the information concealed in the most recent observation, and second the model selection which was used meant that a specification with a trend term was, to some extent unavoidably, only chosen in a few cases.

2.5. Replacement demand per occupation

Because the forecasts in the information system are primarily intended to delineate the labour market prospects of school-leavers, the replacement demand plays an important role, along with changes in employment levels. As was noted in section 2.2, it would be expected that the people who are already working would generally be able to hold on to their jobs or, given their experience, would be preferred for new jobs. The opportunities for the new entrants are therefore limited to the net number of new jobs plus the vacancies created when other people leave the labour market. What is important here is not the flow of people out of the labour market in itself, but the replacement demand — the number of jobs which are made available by people leaving.[8]

Because this replacement demand may account for a very large proportion of the total number of job openings, in the forecasts both the expansion demand and the replacement demand are incorporated as components of the demand for new entrants. At that time the replacement demand was predicted in a relatively simple way. The structure which was used at that time has since been systematically elaborated and improved (Willems and De Grip, 1990 and De Grip and Willems, 1992). The Bureau of Labor Statistics (BLS, 1992) now also employs forecasts of replacement demand. OSA (1993) also maintained that the replacement demand is a very important component

8. This definition of replacement demand is not always used. Sometimes (for example by the CBS, 1953) the replacement demand is supposed to be equal to the outflow.

when compiling a forecast of the future labour market situation. The authors attribute this importance to the absolute magnitude of this component. However Borghans (1991) states that this argument is incorrect. Relevance for forecasts is not determined so much by the absolute magnitude of the individual components, but rather by how much the magnitude changes. If the replacement demand was large but reasonably constant, the future labour market prospects would also not be influenced by this figure. However, it can be seen that, at least for the group of teachers studied by Borghans (1991), the variability of the replacement demand is also substantial.

Measuring the replacement demand

In compiling the replacement demand forecasts, the actual replacement demand over a recent period must first be established. The future replacement demand per occupational class can only be sensibly forecast on the basis of such an actual measure. If the replacement demand in an historical period is to be adequately measured, panel data is in principle required. It must be possible to deduce from the data which people changed their occupational class in a particular period. The sets of panel data which are available are however too small to draw any sensible conclusions at the aggregation level that is used for the information system.[9] For the replacement demand forecasts in the information system, therefore, it has been necessary to rely on the figures for the current state of affairs derived from the Employee Census and the EBB.

The forecasts of replacement demand per occupational class were based on a comparison of the age distribution of the people working in each occupational class, as shown in the Employee Censuses of 1979 and 1985. If the population is divided into 5 year cohorts, after an interval of 5 years everyone will have moved up one age group. By aggregating across the cohorts, it is thus possible to establish approximately the proportion of each occupational class who enter and leave that class. A correction is required to allow for the fact that the cohorts are of five years, and Employee Censuses occur at six year intervals, but this does not affect the essential method.

One disadvantage of this method of measurement, however, is that only the net effects for each cohort are registered. If there are people within one cohort who are leaving and entering an occupational class, it is in fact only the balance of these flows which is recorded. This problem is, in part, solved by dividing those already working in an occupational class by age and gender. In other words, the workers are, so far as possible, divided into homogenous groups. In each group the flows should be in only one direction, whether that be people leaving or entering the occupational class. For men this solution is apparently accurate enough, because they often enter or leave a particular occupational class at specific ages. One exception to this is the group of young people, who quite frequently change their occupation in the first years after entering the labour

9. Panel data have an additional disadvantage for measuring occupational mobility: generally much more mobility is observed than has actually occurred, because of changes in the names which individuals give to their occupation and errors in the classification system (see BLS, 1992, and Glebbeek, 1993). If aggregated data is used, these errors are likely to largely cancel each-other out. It would appear more sensible, for these purposes, simply to use cross-sectional data incorporating the retrospective demand for the occupation in the previous year.

market. Problems can also arise in measuring the movements of women, because they often leave the market temporarily and re-enter later. If there are some women in a given age group re-entering an occupational class, and other women of the same age group leave that class, only the net flow will be recorded. This problem could be reduced if the female population could be further divided into one group who are more likely to enter the occupation and another who are more likely to leave, but it would be almost impossible to find characteristics which would adequately distinguish the two groups. An alternative would be to use the information in the EBB survey which derives from the question as to which respondent's most important activity had been in the previous year.

One problem which is encountered when the population is divided into smaller groups, to improve the observation of flows in and out, is that the data which are used becomes less reliable because the cell filling is less adequate. This means that sampling errors begin to dominate the forecasts. It is difficult to say with any precision what the effect of this may be, but if the number of groups is increased any further, this error element must be included in the analysis. There are statistical methods which could capture this increasing tendency to error adequately. For example, one might consider the 'random coefficient' model which is also used for expansion demand. These problems cannot be observed in the empirical evaluation because the actual replacement demand can only be determined using the same method. The empirical evaluation is therefore limited to the forecast aspect of the replacement demand.

The approach described above leads, in the first place, to a measure of outflow within an occupational class. But for new entrants to the labour market, this outflow is only relevant if it leads to new vacancies. Because the expansion demand is also a net quantity, the replacement demand is by definition equal to the outflow where employment levels are rising. If employment falls, there are more people leaving than entering the market. This difference is also, by definition, equal to the (negative) expansion demand. The relationship between replacement demand and the numbers leaving the market is thus fixed for any given level of expansion demand. Thus:

Replacement demand = outflow + MIN{expansion demand, 0}

On the basis of this equation, the number of job openings can be determined in two ways. The first is to begin with the outflow and add the expansion demand:

Job openings = outflow + expansion demand

The second possibility is to begin with the replacement demand and, if employment is growing, add the expansion demand:

Job openings = replacement demand + MAX{ expansion demand, 0}

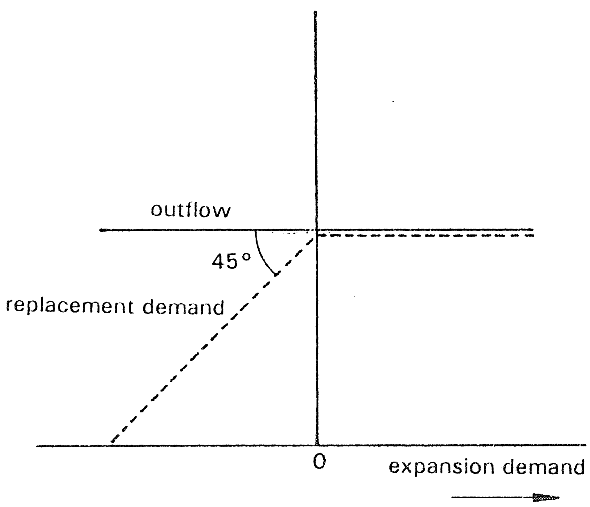
Substituting the definition of replacement demand in the last of these equations shows that the two approaches are equivalent. The definition which begins with replacement demand also shows that the number of job openings can never be negative.

Predicting replacement demand

Because the two definitions are equivalent, it is of no great consequence how the number of job openings is calculated. The current approach, in which the replacement demand and the positive component of the expansion demand are added together, is perhaps more intuitive, because it is easier to consider that the number of job openings is low because jobs which fall vacant are not filled than to say that the job openings which might have arisen due to outflows have been cancelled out by a fall in employment.

In compiling forecasts, however, the distinction between these two ways of conceptualising job openings is important. In the replacement demand forecasts which are being evaluated here, the number of job openings is determined on the basis of outflows. As a consequence, an extrapolation of this outflow, i.e., the outflow for the period 1985-1992 is calculated on the basis of the outflow coefficients observed in the period 1979-1985. This ignores the fact that the number of people who leave the market depends on changes in employment levels. The only correction which has been made is a correction for each age group to allow for changes in total employment levels. Later another correction was incorporated, for changes in the overall rate of participation. Specific changes in the employment for particular occupations, however, have not been considered.

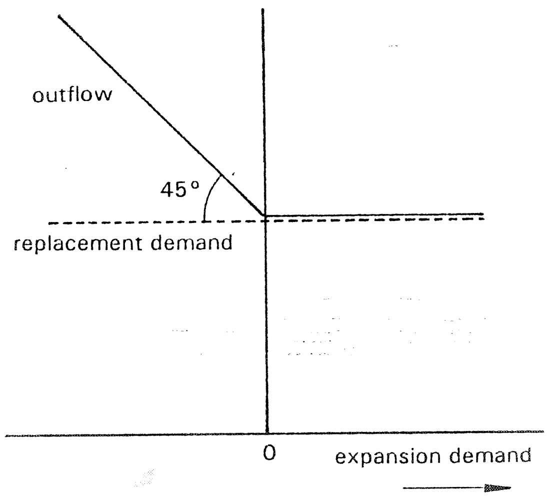
Figure 2.1. Assumed relation between expansion demand and outflow or replacement demand using a forecasting method based on outflows



This approach can be seen from two points of view. The first is to defend it by assuming that occupation-specific changes in employment levels during the forecast period will be the same as those in the period covered by observations. But this supposition is often invalid, and it is inconsistent with the forecasts of the expansion demand, for which changes in the growth in employment are allowed for. A second assumption which could justify the approach described above is that the outflow is independent of changes in employment levels. A reduction in employment in an occupational class would then always be entirely reflected in reduced numbers entering that class. But it is not at all clear that this supposition can be justified. If employment in an occupational class is falling rapidly, to the extent that the reduction in employment is larger than the normal numbers of people entering, such a supposition is actually impossible. Thus if one bases the replacement demand forecasts on figures for employment outflows, there is an implicit assumption regarding the relation between outflows and expansion demand. This relationship is shown in figure 2.1.

On the basis of work by Willems and De Grip (1990), the structure of the replacement demand forecasts was changed from 1990 onwards. Since then, job openings have been defined as the sum of the replacement demand and the positive component of the expansion demand, and the forecasting method has been adjusted to match this definition. The outflow coefficients are however still used to calculate the effect of the age structure of each occupational class on the number of people leaving that class, but the replacement demand is determined on the basis of this forecast assuming that the expansion demand in the forecast period will be equal to the expansion demand in the period covered by observations. This approach boils down, in principle, to extrapolating the replacement demand, but has the advantage that it is not necessary to determine which age group is affected by the changes in employment levels. It is thus assumed that the replacement demand is not dependent on changes in employment levels, which means that the outflow must be dependent on these changes. The whole effect of falling employment is thus manifest in the departure, from the occupational class, of those already working in it, and not in any reduction in the numbers of new entrants. The link between replacement demand, outflow and expansion demand which is assumed here is shown in figure 2.2.

Figure 2.2. Assumed relation between expansion demand and outflow or replacement demand using a forecasting method based on replacement demand



For both of the methods described above, a specific assumption is made as regards the relation between expansion demand and outflow (and thus between expansion and replacement demand). An incorrect specification of this relation would only be relevant for forecasting purposes if the changes in employment levels in a particular occupational class in the forecast period should change radically, as compared to the changes in employment levels in the period covered by the observations. Forecasting errors will arise especially in occupational classes suffering such radical changes. It would be expected that there would in fact be a relation between expansion demand and the numbers leaving the occupational class. When employment is falling, businesses will try to reduce the size of their workforce. Their choice between older employees and new entrants will presumably be related to the relationship between wages and productivity for the two groups. There may also be an aggregation effect: where particular businesses go bankrupt and, at the same time, new businesses are being established, older employees are indirectly making way for new entrants.

Conclusions

All in all, three important problems may be encountered in determining the replacement demand per occupational class. First, the nature of the data means that the numbers entering and leaving each class must be determined on the basis of figures referring to the current situation, so that only the net flows method can be used. To reduce measurement errors which arise with this net method, the workforce must be subdivided, in some adequate way, into groups which experience either only inflows, or only outflows. When this method of measurement has been developed to the point that it is more suited to gauging the flows of older people entering a class, the potential competition between these age groups (people re-entering in the workforce or switching occupations) and school-leavers must be explicitly included in the analysis.

Another related problem concerns the quality of the data which was used. Because the workforce is divided into many small groups, there is a risk that the sampling errors may dominate the forecasts. It would be possible to examine the relation between the number of groups which are defined and the measurement errors. Statistical methods could also be developed to reduce these errors. The forecast of the replacement demand would also improve if it were based on observations from more than one period.

The third consideration is that it is important to obtain more insight into the responses of businesses to various possible changes in the employment level for an occupational group. The question is whether falling demand leads to a reduction in the flows of new entrants or an increase in the number of older employees leaving the occupation. Bad specification of this relationship would be expected to have negative consequences for the forecasts, particularly when the direction of change in employment levels is reversed. It would therefore be useful to carry out more research into the shape of the outflow curve.

2.6. Expansion demand per type of education

As was said in section 2.4, in the expansion demand model the demand for labour is first formulated in the form of a breakdown by occupational classes. The demand from employers, in fact, arises primarily from the need to have some particular tasks, or functions, carried out. But the supply of labour is characterized especially by the type of education which potential workers have completed. If a confrontation of demand and supply is to be made, the forecasts of demand by occupational classes must be transposed into forecasts by type of education, or *vice versa*, the forecast of supply may be translated into the supply per occupational class. Because the primary goal of the information system is to provide forecasts which will be useful for those choosing a course of study or occupation, a confrontation of the demand and supply for each type of education is more suitable.

Generally one does not find only one particular type of education education which is appropriate for a particular occupation. De Grip and Heijke (1991) show that it is important to recognize that people from several different types of education may be eligible for an occupation. The problem therefore arises, that the demand for labour in a given occupational class may be assigned to several types of education. Only one person, with one particular educational background, can be hired to fill each job. In fact people with a particular type of education are generally in demand in several occupational classes, but are in competition in these classes with workers with different educational backgrounds. One logical way of dealing with this conceptual problem is to predict the educational structure of occupational classes on the basis of their structure in the past. This breakdown need not be fixed — it can be related to explanatory variables. If the explanatory variables are independent of supply factors, the predicted demand can be interpreted as *ex ante* demand, i.e., the demand which would have arisen if there had been no changes in the relation between supply and demand.

As noted in section 2.2, comparing the *ex ante* demand with the supply then gives an indication of the tension between demand and supply. This tension will lead to adjustment processes in the labour market which will worsen the position of school-leavers whose education is in over-supply and improve the situation of school-leavers with an educational background for which there is a shortage. The demand-supply ratio thus gives an indication of the improvements or deteriorations which can be expected in the labour market situation of school-leavers with that particular type of education.

An attempt can also be made to explain the adjustment processes within the model. In that case the supply elements do have an influence on the demand for education of particular types. Such an extended model would appear to give more insight into the functioning of the market, but it is important to understand that the adjustment processes which have already occurred within the model may also have already contributed to a deterioration or improvement in the labour market position. Relative demand and supply may lead someone to switch to a different job: if there is excess supply this will obviously be one which is seen as less attractive, and in the event of short supply it will be a better job. The adjustment in the demand can also come about through changes in the conditions of employment. But because the models which are used bring only the employment itself into the analysis, the positions which school-leavers hold in their jobs are never considered. Therefore the relationship between *ex ante* demand and supply is an important indicator of the labour market prospects of school-leavers.

Since the predicted expansion demand per occupation, *ex ante,* need not be the same as the actual future employment per type of education, means that there is no point in carrying out an empirical evaluation of the forecasts of expansion demand in themselves. Only a comparison with the forecasts of supply can show whether the *ex ante* demand was a good indicator was of the future labour market position of school-leavers with a particular educational background.

The forecasting method which was at that time used is comparable to the approach used for the expansion demand per occupational class (Beekman *et al.*, 1989). The model is in fact built in two stages. First the employment for each level of education is estimated and then this is broken down further into the various types of education. In the model, the educational structure of occupational classes is analyzed for each economic sector. An approach at such a low level of aggregation is no longer possible. The CBS's thresholds for the supply of information make it impossible to use data at this level. However Van Eijs and Borghans (1993) show that the educational structure of occupational classes does not vary so much that a specific analysis for each economic sector is necessary. Thus it is in principle sufficient to analyze the educational structure of occupational classes, without distinguishing between economic sectors.

The explanatory variables which were applied to the educational levels of the workers in the various occupational classes included not only the potential workforce per educational level and constants for the occupation and sector, but also, as before, two indicators of technological development: investments related to the added value per sector and the degree of computerisation per economic sector. All but the first of these explanatory variables refer only to the economic sector level and are therefore not useful if sectoral information is no longer utilized. It is however difficult to find relevant explanatory variables at the level of occupational classes. The most promising way to strengthen the economic dimension of an educational model would seem to be making greater allowance in the model for the mutual competition between types of education (see Borghans and Heijke, 1993, and Van Eijs, 1994), so that substitution process and the tendency of people with one type of education to crowd others out can be included in the analysis.

The inclusion of the potential workforce at each level of education as an explanatory variable in the employment equations means that the model already contains a partial adjustment of the demand to the supply, as was shown earlier. This approach has the disadvantage that one component of the alterations in the labour market position of school-leavers which results from the allocation process on the labour market is not registered for the tension indicator. This indicator does however have the virtue of exercising a clear stabilising influence on the forecasts of labour market prospects. Including this variable ensured that the increasing supply of people with higher education also leads to an increasing forecast of the demand for those with higher education. The fact that the increase in the number of those with higher education will also probably lead to a deterioration in their labour market position is thus partly ignored in the forecast results. From this point of view it would also be an improvement if a forecast of the *ex ante* demand was first compiled, without allowing supply factors to play any role, so that the confrontation of demand and supply remains as clean as possible. It would then be possible to go on to say how the labour market prospects for the various types of education would be expected to manifest themselves.

The second step in determining the expansion demand per type of education relates to the breakdown of the expansion demand for educational levels into the demand for each type of education. The methodology which is used is similar, but with no economic explanatory variables.

The model contains only constants for each occupational class and economic sector and a trend term.

Because the models for the forecasts of expansion demand per type of education are similar, as regards methodology, to the model for occupational classes, similar problems arise. It is again difficult to find good and useable explanatory variables and the specification and estimation give rise to similar problems as regards the stability and reliability of the forecasts. Although there is no model selection process, the parameter estimates frequently produce very unstable results because of the very short time series which was available. Moreover, as before, no account is taken in the specification of the information which is concealed in the difference between the explained and the actual number of workers in the last year for which there are observations. There is also an additional problem which is encountered with the educational model: because the *ex ante* and *ex post* demand may differ, the available data, which relates to the *ex post* demand, cannot simply be used to forecast the *ex ante* demand. This problem is dealt with by Borghans and Heijke (1993), who also take the first steps towards explicitly including the mutual crowding-out and substitution effects between types of education in the forecasts for types of education.

Conclusions

Just as was found in dealing with the expansion demand by occupational class, the model for the expansion demand per type of education suffers problems with the instability of parameter estimates, and the specification chosen in this does not fully use the informative value of the most recent observations. Because it is difficult to find suitable explanatory variables at the level of occupations and types of education, and because mutual competition, substitution and crowding-out can play a large role in the demand for types of education, the better use of the last observations is the most promising area to focus attention on in the further development of the model. The robustness of the model is very important in any further development, because large forecasting errors lead to unusable forecasts as regards the discrepancy between demand and supply.

Because the interpretation of the discrepancy between demand and supply is rather abstract as an indicator of tensions in the labour market, it seems important to also forecast how the market will respond to the situation. The Higher Vocational Education Monitor and RUBS school-leaver surveys are very suitable for this. The description of the labour market position of types of education would then contain not only a general characterization of their future labour market position, on the basis of a comparison of demand and supply, but also information on aspects which may result from this position, such as the degree of under-utilization, whether wages are relatively low etc. By also making forecasts of the demand after the market response (the *ex post* demand) the real employment effects of these discrepancies can then also, finally, be predicted. Predictions of how tension in the labour market will manifest itself will however always be surrounded by more uncertainty than the indicative forecasts of the tension in the labour market itself.

2.7. Replacement demand per type of education

The methodology applied in forecasting the replacement demand per occupational class could not, at that time, be used for the replacement demand per type of education. The necessary data on the workforce as regards type of education, gender, and age categories was not available then. It was also not sensible to transpose the replacement demand forecasts per occupational class into replacement demand per type of education, because then the job-to-job mobility would be incorrectly counted as part of the replacement demand per type of education.

For the types of education, as for the occupations (section 2.5 above), it was in fact the outflow which was forecast, since any negative expansion demand and this outflow were simply added together. In place of the forecasting method which was used for the occupational classes, for the types of education the expected replacement demand was based on the proportion of workers with a particular educational background who were more than 55 years old. This is a fairly simple approach which is based on the supposition that, since inter-occupational mobility can be ignored, people from all educational backgrounds only leave the labour market in significant numbers when they reach the older age-groups. It is also assumed here, as for the replacement demand per occupational class, that the outflow is not dependent on the expansion demand for that particular type of education. However a correction is made on the basis of employment forecasts per age-group for the workforce as a whole, just as was done for the replacement demand per occupational class. Later a correction was also added to allow for changes in the overall rate of participation per age category, but neither of these corrections includes any elements which are specific to particular types of education.

The necessary data regarding the workforce, comparable to the data used to forecast the replacement demand per occupational class, is now available. The replacement demand per type of education is now also determined with the method developed by Willems and De Grip (1990).[10] This means that it is not the outflow but the replacement demand which is predicted. The implicit assumption is that the replacement demand does not depend on the expansion demand, and thus that the outflow does. As with the occupational classes, however, it is not clear whether this supposition is more plausible than the supposition that the outflow is not influenced by changes in employment levels. The assumption is in fact based on the supposition that, where there is negative expansion demand, the effects of the squeeze are felt by older employees and not by new entrants. As regards occupational classes, it is quite plausible that those already working may switch, when employment is falling, to other occupational classes, but for types of education this necessarily means that the extra outflow leaves the labour market. This is a much more sweeping

---

10. Willems and De Grip (1990) did not yet have access to data on type of training x age/gender, so their replacement demand forecasts per type of training had to be made indirectly, on the basis of data on occupational class x age/gender and occupational class x type of training. This data is now available, so that the same method can be used for the replacement demand for training types and for occupational classes.

assumption.[11] However, because a fall in the employment available per type of education is rare, the choice between a constant outflow or a constant replacement demand is in practical terms less relevant. Naturally, replacement demand remains a more natural concept than outflow when interpreting the composition of the job openings.

Conclusions

To sum up, it can be said that similar problems are encountered in dealing with the replacement demand per type of education and the replacement demand per occupational class. If the forecasts no longer depend simply on the number of people above a particular age, but are determined on the basis of historical records of outflows, there is first of all a measurement problem. The relationship between the (*ex ante*) expansion demand and the outflow also needs to be examined. It is not self-evident that the same relationship will be found here as was found with occupational classes.

2.8. Supply per type of education

Because the Information System on Education and the Labour Market is intended to indicate the labour market prospects of students, employment forecasts in themselves are not sufficient. An individual will also want to know how many students will be competing for these new jobs. In the forecasts made at that time it was supposed that this competition takes place between school-leavers with education in the same subject area. In later forecasts the short-term unemployed are also included. Because the labour market does not consist of a number of completely separated segments, school-leavers will in fact also face competition from school-leavers who have studied other disciplines. In Borghans and Heijke (1993) a start has been made on developing ways of taking this mutual competition into account in the forecasts.[12]

In compiling forecasts of the flow of students from the education system, the first element is the 'Skill' forecasts, which are provided by the CPB and cover the numbers of students in day-time education and the number who have ended a course, with or without a diploma. Nowadays similar forecasts are compiled by the Ministry of Education and Science, and are called 'Reference estimates'. However the aggregation level of the Skill data was higher than the aggregation level of the types of education which are distinguished in the information system.

The Student Estimates department (*Taakgroep Studentenramingen*) at that time also compiled Worsa and Rhobos figures for university and Higher Vocational Education respectively. These estimates contained a more detailed breakdown of the number of people leaving the education

---

11. Because the *ex post* demand for training types is not equal to the *ex ante* demand, this relation is in any case more difficult to determine than was the case for occupational classes. The *ex post* demand which is observed is in fact not exogenous, but is itself partly determined by flows into and out of the type of education. An adequate analysis would therefore require that a link be made between the *ex ante* demand and the outflow.

12. The competitive position of types of education is dealt with in Borghans (1992) and Van der Velden and Borghans (1993).

system, but both forecasts were in any case based on the Skill forecast of the CPB. The subdivisions which were used in Worsa and Rhobos did not entirely correspond to the classification of types of education used in the information system. Therefore a number of types of education had to be combined, and some had even to be split up.

For Lower and Intermediate Vocational Education and General Secondary Education, no studies comparable to Rhobos and Worsa were available, so it was necessary to work directly from the Skill forecasts. These figures were broken down into the types of education which are distinguished in the information system on the basis of the educational matrix of 1985. Then the estimates of the numbers graduating with a qualification had to be translated into the numbers leaving the education system, because it is also possible for students to move on to further courses within day-time education.

In addition to school-leavers who gained a qualification, Skill also recognized school-leavers who dropped out of a particular type of education without any diploma. With the aid of the educational matrix, this group was assigned to whatever previous educational course they had obtained a diploma in.

Forecasts are also made for flows into the labour market from non-regular education (apprenticeships, recognized correspondence courses, health care training, etc.). Skill does not contain figures for these types of education, so these forecasts had to be based on additional statistics from the CBS, the CORO (a national umbrella organisation for apprenticeship organs), and from direct contacts with the organisers of particular types of education. Because people taking part in these non-regular courses are already available for the labour market before they complete their study, the number of people who successfully completed such a course was distributed as a negative entry over the kinds of previous regular education they had completed before entering non-regular education. It is thus implicitly assumed that those taking a further course in non-regular education have all recently come from the regular education system.[13] The supply category which the information system is attempting to measure consists only of school-leavers.

The supply forecasts are based on a large number of assumptions for specific components of the flow entering the labour market. This is because the available forecasts do not have the required aggregation level and because, especially as regards non-regular education, the available data is incomplete. In the approach used at that time, the Skill forecasts were translated directly into forecasts for the flows of school-leavers entering the market from each of the types of education defined by ROA. This translation from Skill forecasts into ROA types of education was based on current data from additional statistics. This method produced only figures for each of ROA's types of education for the years covered by the forecast, and no figures were produced for current flows into the labour market. The combination of forecasts and current data also meant that it was, at

13. The re-training of people who were already working at the start of the forecast period leads in principle, according to the replacement demand method which is now used, to positive replacement demand for the training type under which they had been registered before their retraining, and since only net effects are observed, will generally lead to negative replacement demand in the field in which they re-train, provided they succeed in their study.

some points, automatically assumed that particular ratios would remain constant over time. If the data permitted, it would perhaps be better if a series of historical figures at the level of ROA types of education could first be devised. This would then provide the basis for forecasts, with the Reference Estimates of the Ministry of Education and Science serving as a restriction to the forecasts.

The fact that the methodology of the supply forecasts produces forecasts directly, without first building up historical data, in fact implies that it is not possible to evaluate these forecasts empirically. To prepare figures which would be suitable for evaluation, at the aggregation level used in the information system, the data would have to be put through a reconstruction process analogous to the adjustments which have occurred in the forecasts. This would mean that these adjustments themselves could not be evaluated. Therefore the evaluation of the forecasts of supply in chapter 4 has been limited to an evaluation of the Skill forecasts.

2.9. The confrontation of demand and supply

If a picture of the future labour market prospects of types of education is to be drawn up, a confrontation of demand and supply must be made. This was done at that time by means of the labour market indicator, or LMI. The LMI is the quotient of the new entrants' supply and demand. As was said earlier, such a ratio between demand and supply must not be interpreted as an indicator of the number of school-leavers who will be unemployed, or the number of vacancies that will remain unfilled.

The real significance of the confrontation between demand and supply is that it shows what tensions can be expected to appear in the labour market. The labour market will presumably largely adjust in response to these tensions, but this will alter the labour market position of people with the types of education concerned. If there is an over-supply of school-leavers, this does not automatically mean that they will be unemployed. It implies that they would be expected to get less attractive jobs and lower wages, and that they will face a higher risk of unemployment.

In determining the ratio of demand and supply, two factors play an important role. In the first place, it is necessary to determine which groups have been included in calculating the demand and supply. Because the primary goal of the forecasts is educational and vocational guidance, the first principle of the labour market indicator is that it should provide a picture of the specific labour market position of school-leavers. But to give an indication of the labour market position which is relevant for school-leavers it must reflect the ratio between the number of jobs that is in principle available for them, and the total size of the group who compete for these jobs.

In the LMI as it was calculated at that time, the supply consists only of the school-leavers themselves. Later this group was extended to include the short-term unemployed at the beginning of the forecast period, who are assumed to constitute serious competition for the school-leavers. The demand in the analysis is naturally only the number of job openings. It is thus assumed that existing jobs which have not fallen vacant because of the departure of the current employees are not available for the school-leavers. If one wanted to assume that in principle competition can always take place for all jobs, then the total employment would have to be included as an element of both the demand and the supply. It would then no longer be necessary, in considering the demand, to take account of the replacement demand. Instead people leaving the market would have to be subtracted from the supply.

A second important factor in determining the future labour market situation relates to the stability of the indicator. Because both the numerator and in the denominator of this indicator include forecasts of flows in the labour market, the indicator will exhibit a certain instability.[14] Especially where the forecast demand is low, this indicator can have extremely high values. It must be remembered here that this is the quotient of two uncertain factors. To make the forecasts sufficiently accurate, the indicator must therefore be somewhat conservative, which can be achieved in two ways. Either the components can be conservatively predicted, bearing this confrontation in mind, or the definition of the labour market indicator itself may include a conservative element.

However, because conservative forecasts of the demand-supply relationships would mainly mean that the differences between demand and supply should not be estimated as exceptionally large, it is difficult to implement this caution in the separate components. Therefore, since 1990, a definition of the IFL (the current name for the labour market indicator) has been employed in which, as a stabilising element, the total employment for each type of education is incorporated in both the numerator and the denominator. At first glance this looks as if the demand-supply confrontation here is again taking place at the level of the whole labour market. Only the treatment of the outflow differs from a confrontation of the total demand for those already working with the total supply, which was described above. But the change in the definition of the IFL can better be interpreted as an application of the prudence principle. If the labour market is in equilibrium, the demand for new entrants will be equal to the supply of new entrants. This demand or supply will be approximately equal to a certain fraction of the total number of workers with that particular type of education. For example, suppose that the equilibrium demand (D_i^{eq}) and the equilibrium supply (S_i^{eq}) are both equal to 7/40ths of the number of workers (7 is the number of forecast years and 40 is the duration of a working career). This equilibrium demand-supply figure can be interpreted as an extremely cautious forecast of both the demand and the supply. To increase the stability of the labour market indicator, the numerator and denominator of the indicator are composed from a weighted average of the supply forecast and the conservative estimate of supply, on the one hand, and the demand forecast and its conservative estimate on the other hand. The definition then reads:

$$IFL_i = \frac{g_S \hat{S}_i + (1 - g_S) S_i^{eq}}{g_D \hat{D}_i + (1 - g_D) D_i^{eq}}$$

In which g_S and g_D are the weights which are attached to both components, i.e., demand and supply.

---

14. In the forecasts of 1989 this ratio could even be negative, because the total demand for new entrants could be negative. However this is no longer possible in later forecasts because the outflow and replacement demand are consistently defined.

This approach has a stabilising effect on the labour market indicator. The stabilising effect is strongest for those types of education for which both demand and supply differ markedly from the equilibrium demand. Considering the uncertainties in the forecasts, the use of such a conservative definition of the IFL would appear to be an important improvement. In fact the only question is what degree of prudence should be introduced. The present definition implies a ratio of $1 : \frac{40}{7} = 1 : 5.7$, because the total number of working people is added to the demand and supply. The definition thus puts a great deal of weight on the equilibrium demand and supply, and relatively little weight on the actual forecasts. To give these weightings a more substantial foundation, the forecasting errors and their influence on the labour market indicator would have to be analyzed in more detail.

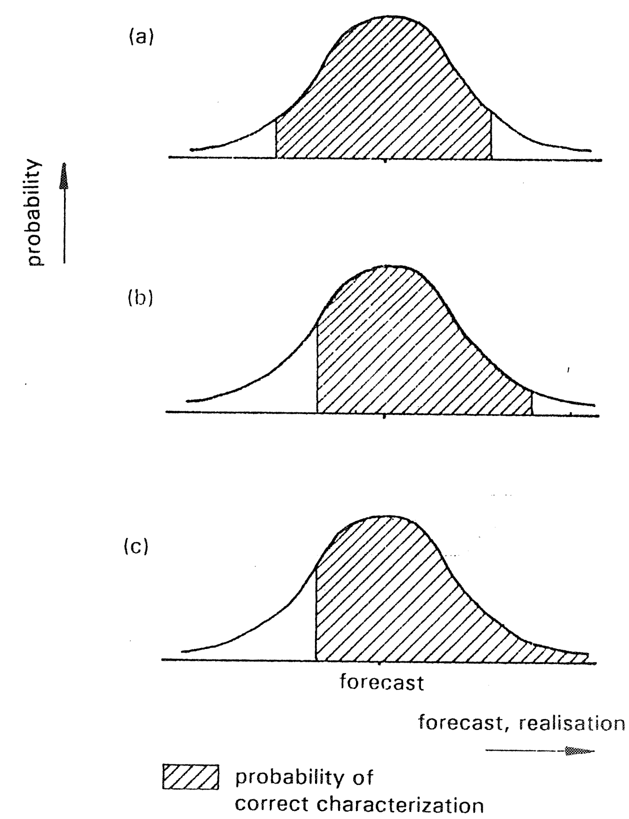
2.10. Qualitative characterizations

The components of the forecasting model which were discussed in the previous paragraphs all relate to forecasting methods which produce point estimates of particular components of the future labour market situation. Such point forecasts suggest a greater degree of precision than can in fact be demonstrated. Moreover, users often find such estimations difficult to interpret unambiguously. Therefore, all quantitative forecast results are translated in the information system into qualitative characterizations on a five-point scale, generally ranging from 'very low', through 'low', 'average', and 'high', to 'very high'. At that time this was done on an *ad hoc* basis, but Wieling, De Grip and Willems (1990) introduced a system whereby the boundaries between the various qualitative characterizations are made dependent on the distribution of the forecast results. This ensures that the forecasts are spread evenly over the characterizations.

This translation into qualitative characterizations clarifies the forecasts greatly, so that this use does not in itself need to be discussed. But the question remains of how the boundaries between the various intervals in the scale should be determined. A number of considerations play a role in this decision. On the one hand it is important to choose a method of marking out the qualitative characterizations in such a way that the chance of getting an appropriate characterization is acceptably high. On the other hand the qualitative characterization has to remain informative. The intervals chosen must not be so large that the characterization has no practical value. Given the size of an interval, it is in principle best to select a confidence interval which extends symmetrically on either side of the estimate in point terms. If it is assumed that the chances of making an underestimate or overestimate are equal, then, for a given interval size, a symmetrical interval gives maximum reliability. But choosing symmetrical intervals would in fact mean that every forecast would be assigned its own interval. It would then not be possible to give these intervals general characterizations such as 'high' or 'average'.

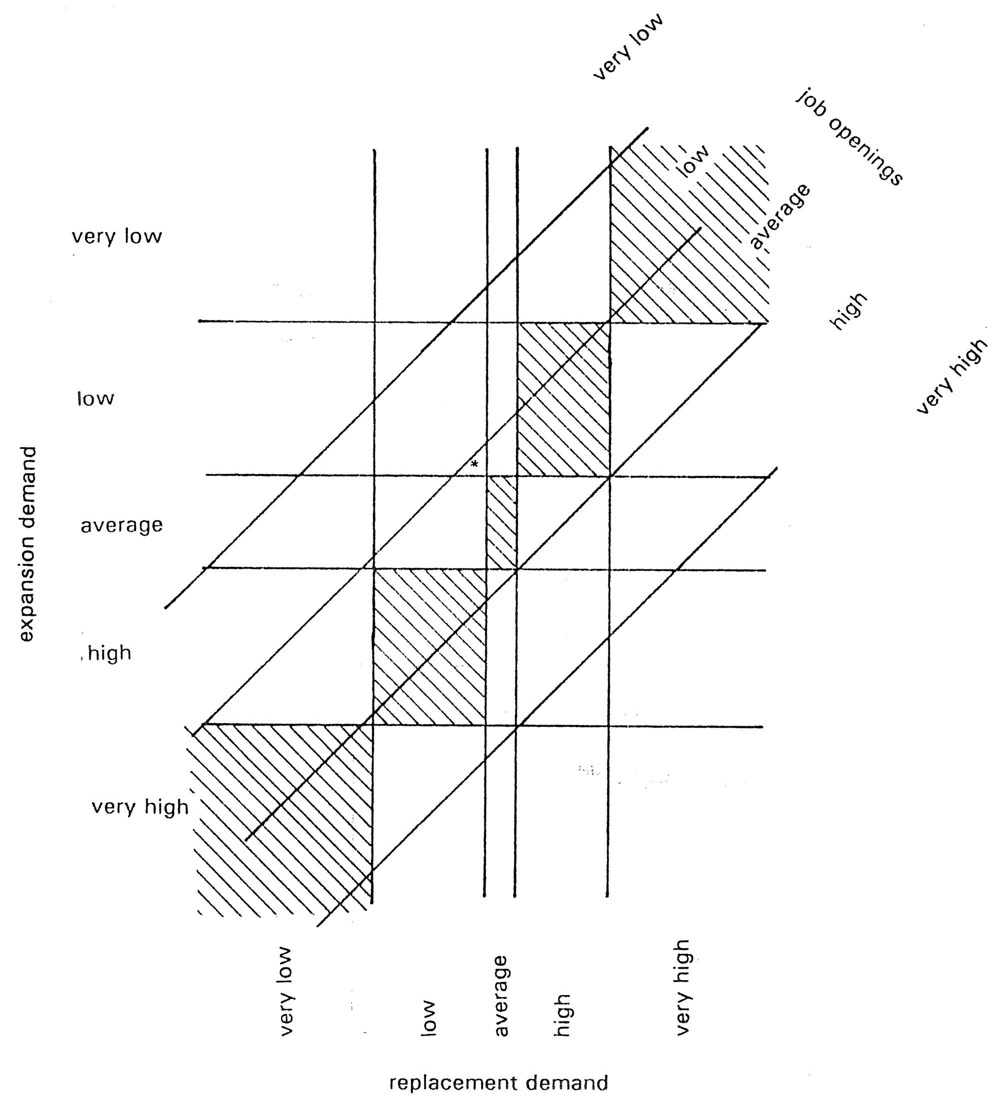
For this reason, to obtain clear qualitative characterizations, the borders between the intervals must be fixed. The consequence is that particular forecasts will fall close to the border between intervals. Figure 2.3 shows that this leads to a reduction in the reliability of the characterizations which are assigned. The figure gives a probability distribution for the realisation of a particular quantity around the forecast. The shaded area in (a) reflects the probability that, if the interval lies symmetrically on either side of the point estimate, the actual outcome will fall within the interval boundaries. This probability will be considerably reduced if the forecast falls close to one of the boundaries of the interval, as in (b). In fact, this is the price, in terms of confidence, of the use of intervals with fixed boundaries. However the qualitative characterization also has two extreme intervals 'very low' and 'very high', which have a boundary on only one side. As curve (c) shows, this will generally very much increase the probability that the characterization will prove to be correct. But if many forecasts fell in these extreme categories, the qualitative characterizations would lose their power to differentiate. Care must therefore be taken to ensure that the only forecasts which fall in these intervals are so extreme that their relative magnitude is not relevant in providing educational and vocational guidance.

Figure 2.3. Probabilities of a characterization using (a) symmetrical intervals, (b) fixed intervals and (c) fixed interval with no boundary on one side



A second important point as regards the determination of the boundaries of the intervals relates to the width of the intervals. Although a characterization should not be directly interpreted as a confidence interval, its size should nevertheless be such as to provide a sufficiently high probability that the characterization will prove to be correct. The actual distribution of the forecasting error thus determines at least the lower limit for the size of the intervals. However, in several of the divisions which were chosen in compiling the forecasts being evaluated here, the interval width chosen for the characterization 'average' was very small, while the characterizations 'low' and 'high' have wider intervals. This approach has also been adopted by Wieling, De Grip and Willems (1990), who justify it by saying that a large number of forecasts are concentrated around the average. The relatively narrow interval for 'average' ensures that the characterizations are spread evenly over the intervals. However one consequence of this is that the reliability of this middle interval is appreciably lower than the reliability of the other intervals. In fact there is no *a priori* reason why characterizations which are close to the average should have smaller forecasting errors than the other characterizations.

Figure 2.4. Relationship between qualitative characterizations of expansion demand, replacement demand and job openings per occupational class



A third consideration in establishing the boundaries of the characterization intervals relates to the harmony between the diverse components of the information system. The job openings for each occupational class are the sum of the expansion and replacement demand. Figure 2.4 shows the interval boundaries of both the expansion and replacement demand and the total number of job openings, and also illustrates their mutual dependence. If replacement demand and expansion demand are given, these then determine the job openings, which lie on the diagonals of the figure.

The interval of the characterization 'average' for the number of job openings is very broad in comparison with the interval width of the characterization 'average' for expansion demand, and even more so if it is compared with the width of the characterization 'average' for replacement demand. This means that an occupational class facing both low replacement demand and low expansion demand may nevertheless have a number of job openings which is characterized as 'average' (see the * in the figure). It would also be expected that high expansion demand and low replacement demand, or a very low expansion demand a very high replacement demand, would cancel each-other out. These areas are shaded in the figure. However it can be seen that, where the expansion demand is very high and replacement demand is very low, there will usually be a large number of job openings, while in the converse situation there will be few job openings.

It should be possible to avoid this sort of incongruity by choosing equal interval widths for each of the separate components, and by keeping the interval widths constant over the entire scale. If the forecasts for one or both of the components of the total number of job openings exhibit a degree of uncertainty which calls for wider intervals, is might be best, in the light of these systematic consideration, to characterize them using a double designation, such as 'low to very low', 'average to low', 'average to high' or 'high to very high'. This maintains consistency with other characterizations, while doubling the interval width. Moreover, these overlapping characterizations would also ensure that the forecasts fall reasonably close to the centre of their interval.

2.11. Risk indicators

In addition to the forecasts of the expected labour market prospects of types of education and the job openings per occupational class, the information system also contains risk indicators for the various types of education and occupational classes, along with current data. The current data is intended to bring the current labour market situation into the picture, in part so that the predictions of the future labour market situation can be seen in the light of the current labour market situation. The risk indicators supplement this information with more structural data on the labour market position of types of education and occupational classes. They provide information on the risk that the labour market situation, as this now appears from the current data and the forecasts for the near future, may change rapidly. They thus give an indication of the degree to which students, once they have entered the labour market, can expect further changes in their market prospects. Moreover, the risk indicators give an indication of the likelihood of unexpected influences altering these prospects even before the forecast period has elapsed. The indicators thus reveal both the long-term prospects and the reliability of the forecasts.

Sensitivity indicator

The sensitivity indicator relates to the sensitivity of the employment levels in the various occupational classes to the state of the business cycle. First, for each economic sector, the extent to which the employment is subject to fluctuations according to the business cycle. The sensitivity indicator is then calculated for every occupational class as a weighted average of the sectoral fluctuations in the changes in employment levels. By relating the sensitivity indicator for occupational classes to fluctuations in the employment in economic sectors, the indicator can be based on a sufficiently long time series. In fact it is not possible to directly measure the sensitivity of occupational classes to the state of the business cycle on the basis of the Employee Census over four years. One disadvantage of this approach is that there is an implicit assumption that all occupations within an economic sector exhibit the same degree of sensitivity to the state of the business cycle.[15] For the economic sectors in manufacturing and industry, it is in fact likely that the shop-floor occupations will be more sensitive to the state of the business cycle than staff positions. Because people with the typical overhead occupations are generally employed in more economic sectors than the more sector-specific production personnel, in practice the sensitivity of the overhead occupations to the state of the business cycle will work out lower than that of an occupational class which is mainly employed within a particular industrial sector.

Switching opportunities

The second risk indicator which is calculated for both the types of education and the occupational classes, indicates the opportunities for switching within the labour market. For occupational classes, the indicator shows how widely their employment is distributed over the economic sectors. For the types of education it relates to the dispersion over both economic sectors and occupational classes. The idea behind this indicator is that if people with a particular education, for example, can work in several occupational classes, they are not entirely dependent on changes in employment levels in one particular occupation, so that they can switch to other occupational classes if the prospects in one occupation suddenly deteriorate.

One shortcoming of this measure is the assumption that the switching opportunities conform to the occupational classes or economic sectors in which the workers in the relevant occupation or with the relevant educational background are at that time working. For types of education which have never suffered the necessity of switching to other occupational classes, the actual dispersion of people over the occupational classes can therefore underestimate the switching opportunities. Therefore a modified switching indicator has since been created in which there is a correction for forced switching to jobs for which workers are over-educated.

---

15. The need for such an assumption was also the reason why this sensitivity indicator is not calculated for each type of training.

3. A METHOD FOR THE EMPIRICAL EVALUATION

3.1. Introduction

The previous chapter discussed the structure of the information system. The forecasts made at that time for the components of this system will be subject, so far as possible, to an empirical evaluation in the following chapter. To give a systematic character to this empirical evaluation, this chapter first describes an evaluation method, beginning in section 3.2 with the problems encountered in the empirical evaluation. Then, on the basis of the objective of the information system, a criterion is introduced against which the system can be evaluated (section 3.3). This criterion makes it possible primarily to evaluate the loss which arises through forecasting errors. To get a better idea of the causes of the forecasting errors, a number of tests are discussed in section 3.4. Finally, section 3.5 deals with the evaluation of the qualitative characterizations.

3.2. Problems in the empirical evaluation

An empirical evaluation is based primarily on a comparison of forecasts with the actual events. Two sorts of problems emerge in evaluating forecasts. The first relates to the forecasts themselves. Naturally the forecasts are always available for evaluation, but sometimes they relate to theoretical concepts which are not even detectable, such as the tension between demand and supply which was mentioned above. This is an indicator of the future labour market situation, which is in general not measurable. The indicator is only meaningful in relation to the possible consequences of expected labour market developments, such as lower wages, higher unemployment, under-utilization, etc. Thus to be able to evaluate the predictions of the tension between demand and supply, the relationship between the tension indicator and quantities which are directly measurable must be clear.

A second possible problem in comparing the forecast and the actuality is the lack of data as regards what actually occurred. This problem arises very frequently and has three causes. First, between 1989, the year in which the forecasts were made, and 1992, the last year covered by the forecasts, a number of classifications were changed. Second, there are a number of factors for which figures were at that time available, for the period to 1988, but which are now not available (for 1992). Third, it sometimes happens that forecasts have been made for quantities for which, even at that time, no figures were known, such as the number of school-leavers who enter the labour market or the replacement demand differentiated by occupation and education.

The two data problems which affect almost every component of the evaluation are the changes which have been introduced in the classification systems and the changeover from the Employee Census to the EBB. As noted in section 2.3, since the first forecasts were compiled in 1989 a number of radical changes have been made in the classification systems which are used within the information system. Current data is collected using the classifications which apply now. Because the CBS does not permit any use to be made of the cells in its data matrixes which contain less than 2,500 people (for annual figures) it is not efficient to gather the data at a lower aggregation level than is necessary for the analysis. However, it proved possible to extract a large part of the required data from the corresponding figures in classification systems which were in use at that time. This flexibility in the classification systems which are used is very important to ensure the continuity of the evaluations. When changing over to a new classification system, it is important, if it is to remain possible to evaluate the forecasts, that the survey data from the past can be transposed into the new classification, and equally that future surveys can be reduced to the old classifications, at least for some years. Such a problem may, for example, also arise with the new occupational classification which the CBS will soon be introducing.

A second frequently recurring data problem in the empirical evaluation relates to the changeover from the Employee Census to the EBB. As was said in section 2.3, this changeover has meant, among other changes, that many more people who work short hours are now counted as part of the workforce. Therefore an explicit lower limit for the number of working hours has been applied in more recent forecasts to remove these very small jobs from the analysis. Because the 1989 forecasts were in principle predictions covering all working people, this criterion has also been used in the evaluation, but the effect is that the working population was considerably larger because of the more inclusive new definition. Most forecasts therefore show a fairly large degree of underestimation. In the evaluation, this problem has been solved by examining relative forecasting errors i.e., the forecasts have been corrected for the forecasting error as regards the overall level of the employment.

An attendant problem was that since 1991, jobs below the 12 hours limit in the EBB surveys have not been encoded. In order to get the best possible figures, despite this difficulty, on the basis of the 1 hour limit, the actual figures have been built up by multiplying the EBB data from 1992, which is based on the 12 hour limit, by the ratio between the total number of jobs and the number of jobs above the 12 hour limit, in 1990. To ensure that this reconstruction was as accurate as possible, this recalculation was not based on the totals per economic sub-sector, occupational class, or type of education, but rather on the separate cells of the underlying matrix, i.e., economic sector x occupational class, or occupational class x type of education.

In addition to these two data problems, which recurred throughout the evaluation, there were other aspects in each component which made a direct evaluation difficult. For the replacement demand, the problem was that, using the EBB data, the flows in the labour market cannot be directly measured. The data which are used in the evaluation are therefore, as in the forecasts themselves, dependent on a part of the model which was used. Thus no empirical evaluation is possible for this component of the model.

The problem in evaluating the supply forecasts was that data on school-leavers is not available in a form corresponding to the classification system used by ROA. In making the forecasts of supply, predictions on the basis of other classifications are translated into the information system classification. Thus there is no data as regards the actual developments. Moreover the translation is based on a number of assumptions regarding the flows coming from non-regular education. Because of the complexity of the translation — which would involve processing information on non-regular education from various sources — and because of the conceptual problem described above, no such reconstruction has been attempted here. Therefore only the Skill forecast of the number of people graduating with qualifications, which is the most important source of the ROA forecasts, is evaluated in this report. However it is important, both for the development of new forecasting models and for the future evaluations, to obtain data on school-leavers which is based on the classification systems in the information system.

3.3. An evaluation criterion

There is often not much point in considering every forecast separately, because many forecasting errors are quite accidental. It is difficult to derive any lessons from these. Therefore it is better to examine the pattern in the forecasting errors. Because ROA's information system includes forecasts for 79 occupational classes and 53 types of education, it is certainly possible and, in view of the quantity of information also desirable, to aggregate separate forecasting errors. This aggregation makes it possible to appraise the probability distribution of the forecasting errors, rather than evaluating every prediction on its own.

This raises the question of how the forecasting errors should be aggregated. A criterion needs to be found by which the separate forecasting errors can be weighted, so that they can be reduced to one measure. This section will discuss the choice of this evaluation criterion and how this measure will be used in this evaluation. In addition to finding a suitable way of combining the separate forecasting errors, judgement must be passed on the quality of the forecast.

As has been said above, the basic principle in assessing the forecasts of the Information System on Education and the Labour Market is the purpose of the forecasts, that they must be suitable for educational and vocational guidance. Most predictions are formulated in terms of people who are working. For an individual student, however, it is not the total number of working people in an occupation or with a given education, or the absolute difference in these numbers, which is interesting. It is mainly a question of the relative difference: $(x_i - \hat{x}_i)/x_i$, in which x_i represents the outcome for some quantity relating to occupation or education i, and $\hat{x}_i$ represents the forecast for the same quantity. Assuming that this relative measure is normally distributed, it is meaningful to interpret the square of the relative variation as the *loss* for each forecast (Granger and Newbold, 1986, pp. 277).

$$(1) \quad L_i(\hat{x}_i) = \left(\frac{x_i - \hat{x}_i}{x_i}\right)^2$$

The loss shows how great the 'damage' resulting from a particular forecasting error is, so that it is possible to calculate the average loss and thus to aggregate the forecasting errors.[16]

16. If the critical level of the expected labour market prospects at which a student would alter his choice was known, a loss function with a specified asymmetric form would have to be selected. Because this information is not known, and because an evaluation does not focus on a particular student but rather on the usefulness of the information for the average student, the quadratic loss criterion also meets the requirements of being able to show how great the damage due to an error has been (see Borghans, 1993, chapter 7).

In aggregating this criterion, it must be remembered that many more students find places in a large occupational group or a large educational category than in a small one. The *average loss is* determined by weighting the separate losses according to the size of the occupational or educational category. This is the evaluation criterion at the aggregated level:

$$(2) \quad AL(\hat{x}) = \sum_i \frac{x_i}{x^{tot}} L_i(\hat{x}_i) = \sum_i \frac{x_i}{x^{tot}} \left(\frac{x_i - \hat{x}_i}{x_i} \right)^2 = \sum_i \frac{1}{x^{tot}} \frac{(x_i - \hat{x}_i)^2}{x_i}$$

Where $x^{tot} = \sum_i x_i$ is the actual total number of workers. The individual loss criterion L_i thus gives an approximation of the loss which is relevant for the choice of an individual student considering a particular occupation or educational course, and the average loss (*AL*) indicates the loss of an average student. The evaluation criteria described above are applicable to the evaluation of forecasts of employment levels. Only a subgroup of the total employment is included in the forecasts of the replacement demand and the job openings. In that case the numerator in the loss function remains the total number of people working, so that the forecasting error remains related to the size of the occupational class or the type of education. These total figures are also used as weights in determining the average loss.

The loss function in (2) is a widely-used criterion, and has also been applied in this context by Bosworth, Evans and Lindley (1974), Evans and Lindley (1973), Borghans and Heijke (1993) and Van Eijs and Borghans (1993), among others, for the evaluation of manpower forecasting methods. The criterion gives an estimate of the distribution of the predictions around the outcome, and so meets the second requirement which Granger and Newbold (1986) propose for an objective evaluation. However this figure does not permit any judgement to made as to the quality of the forecasts. In fact there is no information available as to what a reasonable value for the average loss would be. Such a standard of judgement can be created by comparing the forecast with the predictive quality of other forecasts available at that time, in accordance with the first requirement proposed by Granger and Newbold. A score is defined as the ratio between the predictive quality of the forecast and the predictive quality of a reference forecast. This score is smaller than 1 if the forecast is better than the reference forecast and is more than 1 if the forecast is inferior to the reference:

$$(3) \quad S(\hat{x}, x^{ref}) = \frac{AL(\hat{x})}{AL(x^{ref})}$$

To fill out this definition in practice, a particular forecast must be designated as the reference forecast. In this report the principle is that the forecasts should be compared with the situation which would have faced students if no forecasts from the information system had been available to them. The assumption is that students would then have based their choice on the current labour market situation. Therefore the situation in the base year is used as a reference forecast. This *Same As Before* (SAB) forecast supposes that there will be no changes between the base year and the forecast year. But in some cases, such as in relation to the replacement demand, the Same As Before forecast is not meaningful as a reference point. It is not really plausible to suppose that students know how high the replacement demand is at that moment. In this case it is more plausible to take the average predicted replacement demand as a percentage of the number of working people as a reference forecast for the replacement demand per occupation or type of education.

A last point to be considered is the scale on which the forecasts should be evaluated. When making educational and vocational choices, it is mainly important to get a good estimate of the relative position of an occupation or type of education, and the absolute magnitude of a particular figure is not of prime importance. Therefore the evaluation in this report is generally performed on both the *absolute forecasts* and also on a forecast which has been corrected for the total volume effect (i.e., the forecasts are multiplied by a factor such that the total magnitude of all the occupations (or types of education) is equal to the actual outcome. These forecasts have been called the *relative forecasts*.

3.4. The causes of forecasting errors

The average loss and the score, which were discussed in the previous section, provide information on the quality of the predictions. They do not in themselves show why a forecast has proved to be good or bad, so that it is not clear how the forecasting method could be modified on the basis of the evaluation results. But this is, according to Granger and Newbold (1986), the third important point in making an objective evaluation. To enable conclusions to be drawn on the basis of this evaluation which explain the nature of the problems which can be expected in the forecasts, and to be able to describe the consequences of the methodology which was used, the evaluation yardstick will be supplemented with a number of tests to provide information as to the causes of the forecasting errors.

Concentration of forecasting errors

The evaluation yardstick described in the previous section shows the average forecasting error for all occupations or all types of education. To understand the causes of these forecasting errors it will be necessary to examine the forecasting error in particular sorts of economic sectors, occupational classes or types of education. There are three possible ways of grouping forecasts for such an approach. One might examine the quality of the forecast for each type of education or for each occupation, separately. One disadvantage of this approach is that the element of chance in forecasting errors will not be shown up. Every forecasting error is seen as a unique incident, so that the structure behind the forecasting errors is lost sight of. The usefulness of this information for an objective evaluation is not great, but, using Granger and Newbold's approach (1986), an overview of all the separate forecasting errors can be useful for a subjective evaluation. In fact the very large discrepancies are revealed, so that the possibility of incidental causes for the error in particular forecasts can be examined. For this reason, tables showing errors for each occupation and type of education are included in this report, although only the individual loss has been calculated and not the score, because the score is extremely sensitive to the quality of the reference forecast: if chance factors produce a good reference forecast in a particular case, the score of the ROA forecast is likely to be bad.

A second way of grouping the occupations or types of education is on the basis of the classification criteria themselves, i.e., on the basis of the nature of the economic sector, the occupational class or the type of education. For example, in evaluating the forecasts for types of education it would be interesting to know whether the forecasting errors are different for each educational level.

A third method is to group forecasts according to the methodologies used for particular occupational classes or types of education. In several cases, a different forecasting method was used for some occupations or types of education. It is very interesting to see whether there are differences in the quality of the forecasts produced with these methods. Of course it must be borne in mind that there will have been reasons for the choice of a particular forecasting method. There were grounds for following one methodology for a particular occupation or type of education and a different methodology for another occupation or type of education. Thus there may be an element of selectivity, which could produce a bias in the evaluations of each methodology.

Explaining the forecasting errors

The difference between the forecasts and the actual outcome can best be treated as the outcome of a stochastic variable. At the moment that the forecast is made it is — at least so far as the forecaster is concerned — not possible to predict how large the forecasting errors will be. For evaluation purposes it is interesting to get some idea of the distribution of this random variable: an estimate will be made in the empirical evaluation in chapter 4, based on the premise that the forecasting errors are normally distributed with an expectation of 0.[17] Thus the only parameter which is not known is the standard deviation of this distribution. To give some insight into the nature of the forecasting errors, this standard deviation is estimated as a function of a number of explanatory variables:

(4) $\epsilon_i = x_i - \hat{x}_i \sim N(0, \sigma_i)$

With

(5) $\sigma_i = (x_i)^{\alpha} e^{Z_i \beta}$

In which x_i is again the amount of the quantity for which the prediction was made (for example occupational class or type of education) in the forecast year (1992) and Z_i are the other explanatory variables for the standard deviation of the forecasts. The first factor in equation (5), the size of the occupation or the educational category, is included to determine the effect of scale. Larger forecasting errors would be expected for the larger occupational classes or types of education, but the increase in these forecasting errors will probably be less than proportional. α will thus be less than 1 and more than 0. The choice of the other explanatory variables for the size of the standard deviation of the forecasting errors depended on the quantity being considered. These explanatory variables can be grouped in categories similar to the classification above, i.e., into variables relating to the nature of the quantity covered by the forecast, and variables connected with the methodology used.

---

17. Borghans (1993, chapter 7) has shown that the supposition that the forecasting errors have an expectation of 0 is unavoidable in any coherent approach to this problem. In fact, for the relative quantities this supposition is true by definition.

As noted above, the Information System on Education and the Labour Market also contains 'risk indicators'. These indicators are partly intended to give an idea of the sensitivity of a particular occupational class or a type of education to exogenous influences. A relation would be expected between these risk indicators and the magnitude of the forecasting error. The method based on equations (4) and (5) can also be used to indirectly evaluate these risk indicators.

The form of (5) was selected to ensure that the standard error is always positive. Equations (4) and (5) are estimated using the maximum likelihood method.

Over-estimation or underestimation of changes

One important component in most forecasts is the trend in the quantity which is to be predicted. In extrapolating existing trends it is crucial to establish how this extrapolation should be effected. Simply extending trends often leads to improbable results, especially in the longer term. In that case the trend is only partially incorporated in the forecast, or an approach is chosen which reduces the influence of the trend over time. In principle, similar problems emerge with every explanatory variable. If a parameter estimate happens to be high, the influence of this variable will continue to be overestimated. Therefore, a parameter for which the estimate was not significant when compiling the forecast is often ignored.

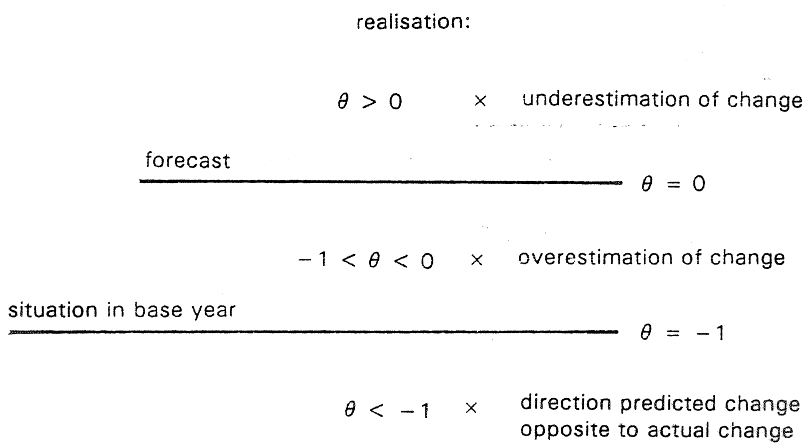
An evaluation is a suitable way to examine the extent to which the estimated influences of the exogenous variables in the past have been properly extended into the future. An extrapolation which was too cautious will appear as an under-estimation of change, while an exaggerated extrapolation will over-estimate the change (see Theil, 1958, pp. 68 and Borghans, 1993, chapter 10). By estimating the following equation it is fairly simple to test for the over or under-estimation of changes:

(6) $x_i - \hat{x}_i = \theta(\hat{x}_i - x_i^{ref}) + \epsilon$

In which x_i^{ref} is the reference forecast described above, in which no changes are expected. On the basis of this equation it is possible to formulate a prediction of the difference between the actual figures and the forecast. The right side of the regression equation (6) contains only quantities which were known at the time the forecast was compiled, so it would in principle have been possible at that time to devise an improved forecast without this over-estimation or under-estimation of changes. Figure 3.1 shows what particular parameter values of θ signify. In the figure it is assumed that the forecast is higher than the value of the variable in the base year, although the reverse is also possible. The situation then would be the mirror image of the situation shown in figure 3.1.

Three areas are marked out in the figure by two limiting values. If $\theta = 0$ then the actual figures are, on average, equal to the forecast. The quality of the forecast may nevertheless be bad, but there is no systematic over or under-estimation of the changes. If $\theta > 0$ then the actual figures for the predicted quantity lie, on average, further from the value in the base year than the forecast, so that the changes have been under-estimated. $\theta < 0$ on the other hand, means that the actual figures were, on average, closer to the value in the base year than the forecast, so the changes have been over-estimated. The assumption here is that $\theta > -1$. If $\theta = -1$ then the over-estimation of changes is total, so that none of the predicted changes had any informative value, and one might equally well have taken either the reference forecast or the ROA forecast. If θ is smaller than -1, there has been an inversion effect. This means that, where falls were predicted there has, on average, been a rise, while quantities for which rises were predicted have, on average, fallen.

Figure 3.1. Possible values of the over-estimation coefficient θ, for over-estimation and under-estimation of changes



An over-estimation of changes need not necessarily have been caused by over-extrapolating trends. Often it results from a large random element in the forecasts. If forecasts contain some noise in addition to their informative qualities, it may be sensible to partly ignore the predicted changes. The larger this random element is, the more cautiously the forecasts should be treated. In (6) this would be expressed by a value for θ of close to -1.

3.5. The evaluation of qualitative characterizations

The evaluation method which has been outlined above has thus far referred only to the point forecasts produced by the forecasting models in the information system. But as has been said, the forecasts are ultimately presented by means of a qualitative characterization of the forecast results.

The most direct way of evaluating these qualitative characterizations is by assigning qualitative characterizations to the actual outcomes, using the same classification system, and then to examine how often this corresponds to the original characterization (see also De Grip, Heijke and Berendsen, 1991). Such an evaluation is made on the basis of a matrix, with the original characterization and the outcome on the two axis. The forecast quality and the method of characterization are then in fact being evaluated simultaneously. Because the previous paragraphs have already described a methodology, based on the point forecasts, for evaluating the forecast quality, in this report the evaluation of the qualitative characterizations will focus mainly on the characterization method.

As was said in section 2.10, the characterization methodology which has been used departs from the confidence intervals which are usually used in statistics, because the boundaries of the qualitative characterizations are independent of the separate forecasts. As has been noted, this leads to two important aberrations. In the first place, the point prediction need not fall in the centre of the confidence interval using this approach, which means that the probability of a correct characterization is smaller. In the second place this characterization system leads to two extreme intervals, on either end of the scale, each of which is bounded on only one side. This means that the chance that a characterization falling in these extreme intervals will be correct is greater because, if the actual outcome is more extreme than the forecast, the forecasting error will not lead to an incorrect characterization.

To determine the reliability of the intervals which were chosen, and to enable the effects of the characterization method to be enumerated, we calculated, for each separate forecast, the probability of the actual outcome falling in the same interval as the forecast. The calculation was carried out in three ways, and was based on the standard deviation of the forecasting errors, which was explained in equations (4) and (5). First, we calculated what the probability of a correct characterization would have been if symmetrical intervals had been employed. This probability is equal to the area under the probability distribution curve in figure 2.4 (a). The size of the intervals used is the same as those used in the information system for the forecast concerned. Because the interval size for the two extreme characterizations ('very low' and 'very high') is unlimited, the interval size in these cases has been supposed to be equal to the size of the neighbouring characterizations ('low' and 'high').

The next step is to examine the probability that the forecast and the actual outcome would fall in the same interval if the boundaries between intervals were randomly placed in relation to the forecast. This probability is not determined on the basis of the standard deviation, but rather on the basis of the actual forecasting error. If the forecast's position within the interval is random, the reciprocal relationship between the forecasting error and the length of the interval determines the probability of a correct characterization. Because this probability is based on the actual forecasting errors and not on the probability distribution, it is not possible to compare it with the probability of a correct characterization for each separate forecast if the interval had been symmetrical. The average probabilities over all forecasts can however show how much the accuracy has been reduced because of the asymmetric intervals. The probability of a correct characterization, measured in this way, is equal to 0 if the forecasting error is greater than the interval length. The number of times that this 'probability 0' occurs therefore gives an indication for the usefulness of the interval length which was chosen.

Finally, the probability that the interval boundaries which were actually chosen have resulted in a correct characterization is determined, again on the basis of the probability distribution of the forecasting error. Both the negative influence of the asymmetry and the positive influence of the unbounded side of the outermost intervals play a role in this actual probability.

4. THE EMPIRICAL EVALUATION

4.1. Introduction

This chapter reports on the empirical evaluation of the forecasts which were made in 1989 for the first version of the Information System on Education and the Labour Market to cover the full width of the education system. The procedure described in the previous chapter has been used in the evaluation wherever possible. As has already been noted, it is not possible to evaluate the forecasts of the information system empirically on certain points. Therefore this chapter contains no evaluation of the expansion demand per type of education and the forecasts of supply per type of education. In addition to ROA's forecasts, the forecasts which served as inputs for the information system are also tested. This chapter therefore contains a section on the forecasts for economic sectors which are compiled by the CPB, and, in the section on the supply, the forecast quality of the Skill forecasts for the education system is examined.

4.2. CPB forecasts for economic sectors

The forecasts of expansion demand are, to a significant extent, based on the CPB's forecasts for economic sectors. Because the quality of these forecasts can largely determine the quality of the forecasts per occupational class and per type of education, the CPB forecasts are scrutinized in this section.

The forecasting error

The forecasts for economic sectors which are used for the expansion demand forecasts are derived from the CPB's estimates of economic sectors, on the basis of the beta model (Eijgenraam and Verkade, 1988). This model is no longer used, and the forecasts for economic sectors are now based on the Athena model. Because in 1988 the CPB did not provide forecasts extending to 1992, the forecasts to 1990 which were available were extrapolated up to and including 1992, based in part on the advice of the CPB.

The CPB predicts the employment per economic sector in terms of labour volumes, based on the data series contained in the *National Accounts* (*Nationale Rekeningen*). Because they are differentiated by type of education and occupational class, the forecasts of expansion demand use data from the Employee Census and (nowadays) the EBB, which relates to the number of people working. Therefore the CPB's labour volume figures must be translated into numbers of people working. This translation is based on CPB forecasts of the ratio between these two quantities, the P/W ratio (Persons/workyears).

However, despite this translation, a difference remains between the data in the *National Accounts* and the Employee Census or EBB. Table 4.1 shows how large these differences were in 1985. In a number of economic sectors the difference between these two data sources is very considerable. In the sector of 'public administration and education', the Employee Census has 14% less people than the *National Accounts*, while for the base metals industry the Employee Census has 15% more people. The differences arise from the use of different definitions and a different classification method.

Table 4.1. Differences between number of people working in 1985 according to the Employee Census and the National Accounts

Occupational sector	Employee Census	National Accounts	difference	%
Agriculture, fishing, forestry	265,000	268,000	−3,000	−1.2
Manufacture of foodstuffs, beverages, tobacco products	165,000	166,000	−1,000	−0.7
Manufacture of textiles, wearing apparel, leather and footwear	55,000	56,000	−1,000	−0.8
Manufacture of wood and building industries	73,000	73,000	0	−0.4
Manufacture of paper and printing and publishing industries	108,000	115,000	−7,000	−5.8
Chemical industry and manufacture of rubber and plastic products	115,000	112,000	3,000	3.0
Base metal industries	39,000	33,000	6,000	15.0
Manufacture of metal products, mechanical and instrument engineering	255,000	254,000	1,000	0.3
Electrical engineering	100,000	101,000	−1,000	−0.5
Manufacture of transport equipment	71,000	71,000	0	0.5
Petroleum industry	10,000	10,000	0	−0.2
Mining and quarrying	11,000	11,000	0	−0.3
Electricity, gas and water	44,000	44,000	0	−0.1
Ownership of dwellings	415,000	417,000	−2,000	−0.4
Trade	702,000	720,000	−18,000	−2.6
Sea and air transport	50,000	51,000	−1,000	−1.2
Transport, storage and communication	267,000	272,000	−5,000	−1.7
Banking and insurances	188,000	186,000	2,000	1.0
Other private services	550,000	561,000	−11,000	−2.0
Medical and veterinary services	374,000	375,000	−1,000	−0.2
Other public services	429,000	427,000	2,000	0.5
Public administration and education	688,000	781,000	−93,000	−13.5

This discrepancy between the data sources must be borne in mind when compiling forecasts, but even after making a correction the difference between the two sets of data remains a potential source of inaccuracy. Therefore, later in this section, we will examine whether this discrepancy had a significant influence on the forecast results.

The forecasts of expansion demand were differentiated for these 22 economic sectors. Table 4.2 gives an overview of the employment per economic sector in 1985, the forecasts for 1992 and the actual outcome in 1992. The table also contains the forecasting error. The individual loss is determined on the basis of this forecasting error.

Table 4.2. Forecasting errors of the CPB-projections of the number of people working per economic sector

Occupational sector	number 1985	forecast 1992	outcome 1992	forecast error	loss
Agriculture, fishing, forestry	265,000	266,000	259,000	7,000	0.0008
Manufacutre of foodstuffs, beverages, tobacco products	165,000	172,000	178,000	−6,000	0.0010
Manufacture of textiles, wearing apparel, leather and footwear	55,000	58,000	50,000	8,000	0.0302
Manufacture of wood and building industries	73,000	79,000	80,000	−1,000	0.0001
Manufacture of paper and printing and publishing industries	108,000	132,000	166,000	−34,000	0.0414
Chemical industry and manufacture of rubber and plastic products	115,000	134,000	144,000	−10,000	0.0048
Base metal industries	39,000	35,000	31,000	4,000	0.0136
Manufacture of metal products, mechanical and instrument engineering	255,000	281,000	322,000	−41,000	0.0162
Electrical engineering	100,000	107,000	115,000	−8,000	0.0054
Manufacture of transport equipment	71,000	73,000	71,000	2,000	0.0005
Petroleum industry	10,000	10,000	10,000	0	0.0001
Mining and quarrying	11,000	12,000	9,000	3,000	0.0944
Electricity, gas and water	44,000	46,000	46,000	0	0.0000
Ownership of dwellings	415,000	448,000	430,000	18,000	0.0017
Trade	702,000	862,000	978,000	−116,000	0.0141
Sea and air transport	50,000	60,000	56,000	4,000	0.0033
Transport, storage and communication	267,000	312,000	350,000	−38,000	0.0121
Banking and insurances	188,000	207,000	229,000	−22,000	0,0087
Other private services	550,000	666,000	908,000	−242,000	0.0709
Medical and veterinary services	374,000	423,000	537,000	−114,000	0.0454
Other public services	429,000	475,000	573,000	−98,000	0.0299
Public administration and education	688,000	809,000	952,000	−143,000	0.0225

The largest forecasting loss was for the mining and quarrying sector. The CPB predicted a growth in the employment in this economic sector of 8%, but it actually fell by 18%. Relatively large losses also occurred for other private services, medical and veterinary services, the paper printing and publishing industry, the textile industry, other public services and public administration and education. In these economic sectors, the growth in employment was much greater in reality than had been predicted. For the other private services, the CPB predicted growth in employment of 21% for the period 1985-1992, while the actual growth was 65%. However this economic sector includes people working through temporary employment bureaus, so it will certainly include many small jobs. In total there were 171,000 jobs in 1992 which lay under the 12-hour threshold. In the medical and veterinary sector the prediction of +13% was also lower than the actual outcome (+44%). For the paper, printing and publishing industry, growth of 22% was forecast, and 54%

growth was achieved. For other public services the forecast was +10% and the actual outcome was +34%. For the public administration and education sector, finally, growth of 18% was predicted, and the actual growth was 39%. There may well be a tendency, as regards the government and semi-governmental sectors, to underestimate the future growth in employment on the basis of the good intentions set out in policy documents. However it is not clear to what extent the high actual growth rates may result from the changeover from the Employee Census to the EBB, or what role the small jobs play in this growth.

The forecast quality

As has already been noted in section 1.2, one must be very cautious in interpreting individual forecasting errors. There are always reasons to be found after the event which explain why the actual outcome differs from the forecast in a particular case. It is therefore more important to compare the forecasts and the actual outcomes as a whole. Table 4.3 gives a picture of this confrontation, comparing the quality of the forecasts which were made, on the basis of the average loss (equation 2) with the Same As Before forecast (SAB). In the first row the two forecasts are directly compared with one another. The average loss of the CPB forecast is considerably lower than the loss of the SAB forecast. This means that the CPB forecast has just 36% of the error which the SAB contains. This would appear to be a reasonably good forecasting result. The quality of both the CPB and SAB forecasts was sharply reduced because the total employment was not correctly forecast. This 'error' is, as has been noted, partly the result of the way in which employment was measured.

For educational and vocational choices, however, it is mainly the relative relationships which are important, since students face a choice between the existing alternatives. Forecasting errors as regards movements in the whole market are, from their point of view, of lesser significance. The second row shows the quality of these relative forecasts. Because the SAB, by definition, has no prediction of growth at all, and the CPB forecast merely under-estimates this growth, the error in the SAB forecast is reduced by more than the error in the CPB forecast. The score of the economic sector forecast, as compared with the SAB, now comes to 47%. However it must be noted that the CPB forecasts are based on actual information up to and including 1988, in contrast to the other components of the information system, for which the most recent available data was for 1985.

Table 4.3 goes on to show the forecasting error, distinguishing between the agricultural and manufacturing sectors on the one hand and the private and public service sectors on the other hand. In absolute terms, the forecasting errors for each of these broad economic areas were roughly equal. This is largely due to the substantial under-estimation of the growth of the service sectors in general. If we consider the relative figures, the forecasting quality (as compared with the SAB forecast) is considerably less satisfactory for the service sectors, while the forecasts for agriculture and manufacturing are somewhat better.

Table 4.3. Overview of the quality of the forecasts by economic sectors

Method	av. loss forecast	av. loss SAB	score
Absolute	0.0254	0.0699	0.36
Relative	0.0120	0.0256	0.47
Agr./manuf., absolute	0.0112	0.0317	0.35
Services, absolute	0.0296	0.0811	0.36
Agr./manuf., relative	0.0191	0.0561	0.34
Services, relative	0.0099	0.0166	0.59

An explanation of the forecasting error

Table 4.4 shows the explanations of the standard deviation in the errors in the forecasts for economic sectors. The first explanatory factor which has been used is a scale variable, showing the size of the economic sector. Additional factors which explain the standard deviation in the forecasting errors are an indicator showing whether the economic sub-sector being considered belongs to the agriculture and manufacturing group, and another variable expressing the relative discrepancy between the figures from the *National Accounts* and the Employee Census. The sector dummy has been included so that structural differences between the two main sectors of the economy can be identified, as was done in table 4.3. The second variable, the difference between the two sources of data, is intended to provide an explanation of the forecasting error arising because forecasts from the *National Accounts* have been used to forecast EBB figures. The greater the differences between these figures, the more difficult it will be to base adequate forecasts on them.

Table 4.4. Explanation of the standard deviation of the forecast errors

Variable	parameter	t-value
Scale	0.62	5.58
Constant	−2.80	1.97
Agr./manufacturing	−0.22	0.66
Diff. Nat.Acc - Employee Census	0.01	0.57

The estimated value of the scale parameter is 0.62. This implies that the increase in the forecasting error is less than proportional to the size of the economic sector. We can also deduce from table 4.4 that the forecasting error in the agricultural and manufacturing sector is indeed smaller, and that the discrepancy between the *National Accounts* and the Employee Census has had the expected influence on the forecasting error. However neither effect is significant, so one must be cautious in drawing any conclusions on this basis.

Under-estimation of changes

The last step of the empirical evaluation of the forecasts for economic sectors is to examine the degree to which changes have been under-estimated. This is shown in table 4.5. Because of the under-estimation of the total changes in employment levels, the distinction between the absolute and relative measures is also important here. In both cases the changes have been significantly under-estimated. The under-estimation coefficient is about 1, meaning that the observed changes are on average about twice as large as was forecast. It would appear that the CPB was rather cautious in compiling its employment forecasts for economic sectors.

Table 4.5. Under-estimation of changes in the forecasts for economic sectors

	under-estimation	t-value
Absolute	1.17	7.36
Relative	1.07	3.86

Figure 4.1 shows the relation between the predicted and the actual change, as compared to the base year, with the forecasts on the horizontal axis and the actual outcome on the vertical axis. If there had been no under-estimation or over-estimation of the changes, the observations would lie around the 45% line. But it can clearly be seen from the figure that the actual growth in employment has been under-estimated, because the changes in the actual outcome are systematically higher than the changes in the forecasts. Figure 4.2 gives the same information for the relative forecasts, i.e., after correcting for forecasting errors regarding the general rate of growth. The relationship is less clear using this relative measure than when using the absolute figures, but a significant under-estimation of changes can still be seen.

Relationship between forecasts for economic sectors and occupational classes

This section has focused on the forecasts for economic sectors. However the quality of this data is only important for the information system in so far as the forecasting errors affect the employment forecasts for occupational classes, and the employment forecasts per type of education which are derived from the occupational forecasts. Table 4.6 gives an overview of the quality of the CPB forecasts, translated into occupational classes, as an aid in assessing the relevance of the errors in the forecasts for economic sectors.

Figure 4.1. Under-estimation of changes in the CPB's forecasts for economic sectors (absolute)

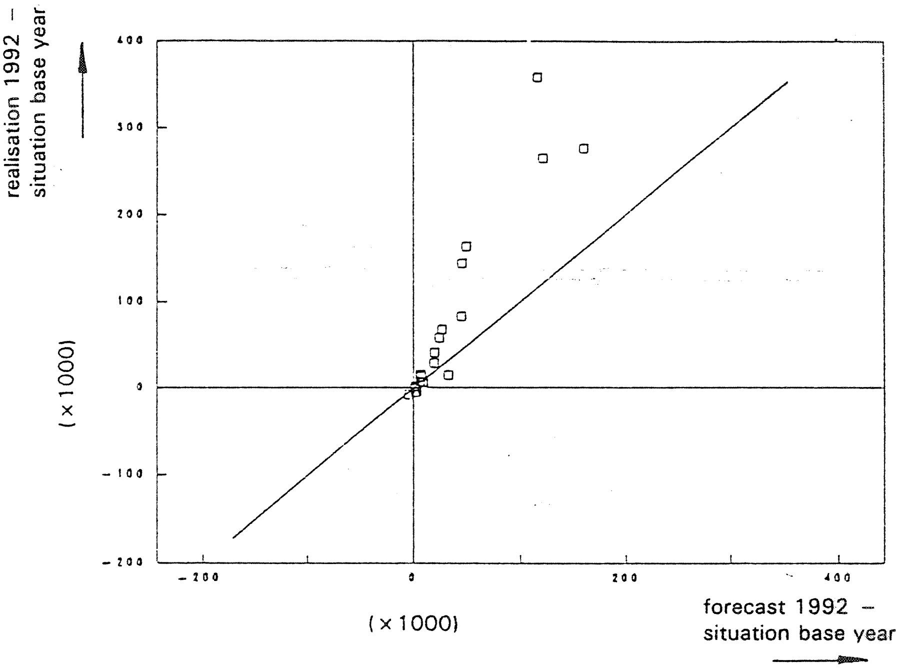
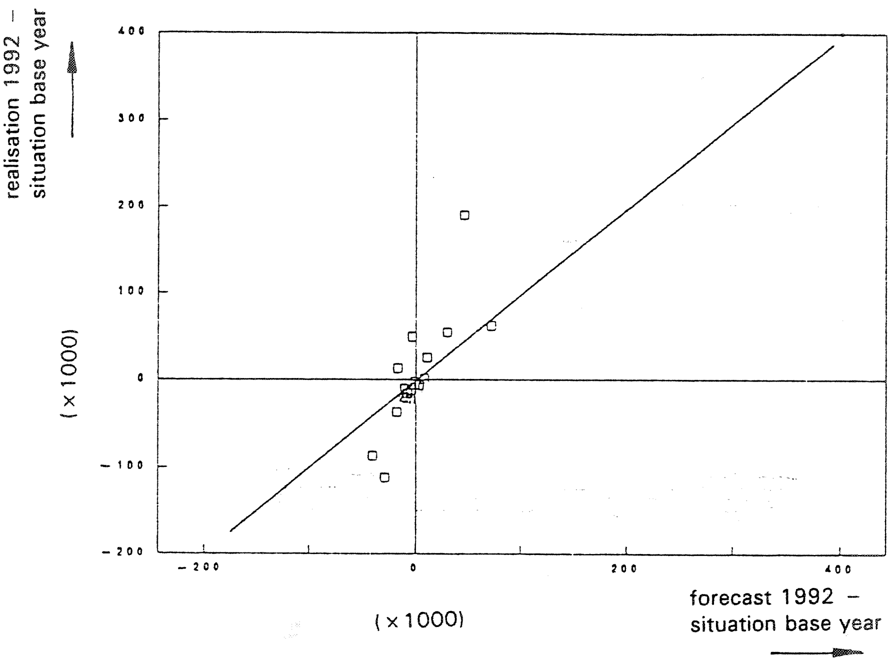


Figure 4.2. Under-estimation of changes in the CPB's forecasts for economic sectors (relative)



To eliminate the influence of the methodology which has been employed for the forecasts for occupational classes, the translation has been based on the fixed coefficient model. The forecasts for economic sectors have been transposed into forecasts for occupational classes using the occupational structure of the economic sectors in both 1985 and 1992. The table gives an overview of the forecast results for both structure matrices, on the basis of the Same As Before forecast, ROA's forecast and the actual employment for each economic sector.

The calculations based on the matrix of the occupational structure for 1985 show that the forecast has approximately the same predictive power as would have been achieved if the actual sector size in 1992 had been used as a starting point. The score, in relation to the SAB forecast, is 76%. This is considerably less good than the 36% from table 4.3, which was found by directly evaluating the forecasts for economic sectors. An examination of the relative loss in the forecasts also shows that the improvement in comparison to the Same As Before forecast is due entirely to a better estimate of the total level of employment. Although the forecasts for economic sectors are, in themselves, good predictions, it would appear that they are not improved by using the structure matrix from 1985.

The second part of the table shows that, if the correct structure matrix is used, the forecast results improve considerably. If the actual employment in the economic sectors for 1992 is used the forecasting error is, by definition, eliminated. If the forecast employment levels are used, the forecasting error is more than halved. This means that, using a fixed coefficient approach, 45% of the forecasting errors regarding occupations derives from the forecasts for economic sectors and 55% of the error is caused by unforseen changes in the occupational structure of the economic sectors (if the SAB forecast is used as a basis, these shares are 73% and 27% respectively). Similar ratios are found if the relative measure is used. This shows that it is very important to take account of possible changes in the occupational structure when drawing up the forecasts of the employment per occupational class.

Table 4.6. Quality of the forecasts by economic sectors, transposed into occupational classes

Economic sector forecast	structure-matrix	average loss (absolute)	average loss (relative)
Base year '85 (SAB)	'85	0.1055	0.0767
Forecast '92	'85	0.0798	0.0771
Outcome '92	'85	0.0806	0.0760
Base year '85 (SAB)	'92	0.0775	0.0518
Forecast '92	'92	0.0358	0.0418
Outcome '92	'92	0	0

The score using the fixed coefficient model is 1.01 using the structure matrix for 1985 and 0.54 if the structure matrix of 1992 is used. Even using the correct occupational structure, the score of the forecasts for economic sectors at the level of occupational classes (absolute 0.0358 ÷ 0.0775 = 0.46, and relative 0.0418 ÷ 0.0518 = 0.81) remains worse than the scores (0.36 absolute and 0.47 relatively) at the level of economic sectors. Apparently a number of comparatively badly forecast economic sectors have had an important effect on the translation of these forecasts in terms of occupational classes.

Conclusions

In brief, the quality of the CPB forecasts which have been used as inputs for the demand side of the model can be said to be good, but the match between the *National Accounts* and the EBB data remains a matter of concern. It would also be expected that the new CBS definition of the workforce, by excluding jobs which are less than 12 hours per week, will improve the quality, because the CPB mainly predicts labour volumes, and the translation into numbers of people working is, especially for the small jobs, very sensitive to the accuracy of the P/W ratio. This group of small jobs is probably also not very relevant for the main target group of the forecasts in the information system, i.e., students who have to make their educational and vocational choices. These small jobs can however be significant for questions regarding people re-entering the workforce. The predictive quality of the forecasts for economic sectors is no lower when they are translated in terms of occupational classes. Using the SAB forecast at occupational level as a basis for comparison, about two thirds of the total forecasting error is caused by movements between economic sectors, and one third by changes in the occupational structure. Of that two thirds of the potential forecasting error, about half disappears if the CPB forecasts for economic sectors are used as a basis. An explanation of the changed occupational structure is therefore indispensable for good forecasts of the employment per occupational class. Finally, it is striking that the forecasts for economic sectors have very much under-estimated changes in employment levels. In later evaluations it will be necessary to see whether this is incidental, or whether changes are being structurally under-estimated.

4.3. Expansion demand per occupation

In this section the accuracy of the forecasts as regards the expansion demand per occupation is empirically tested. Table 4.7 contains an overview of the 79 occupational classes for which a forecast was at that time made. The table shows the number of people working in 1985, the forecast for 1992 and the actual outcome in 1992, in that order, and then gives the forecasting error and the loss per occupation.

Table 4.7. Quality of the forecasts of expansion demand per occupational class[18]

CBS-occupational class	number 1985	forecast 1992	outcome 1992	forecast error	loss
01 Physical scientists and related technicians	34,000	38,000	39,000	−1,000	0,0012
02 Architects, engineers and related technicians	178,000	205,000	240,000	−35,000	0,0207
04 Aircraft and shifs' officers	17,000	19,000	14,000	5,000	0,1374
05 Life scientists and related technicians	29,000	28,000	37,000	−9,000	0,0517
06 Medical, dental, veterinary and related workers	267,000	323,000	385,000	−62,000	0,0265
08 Statisticians, mathematicians and related workers	51,000	96,000	122,000	−26,000	0,0459
09 Economists	15,000	17,000	24,000	−7,000	0,0865
11 Accountants	12,000	15,000	28,000	−13,000	0,2214
12 Jurists	18,000	20,000	32,000	−12,000	0,1387
13 Teachers	275,000	305,000	335,000	−30,000	0,0080
14 Workers in religion	10,000	11,000	11,000	0	0,0042
15 Authors, journalists and related workers	20,000	20,000	39,000	−19,000	0,2319
16 Sculptors, painters, photographers and related creative artists	29,000	30,000	51,000	−21,000	0,1721
17 Composers and performing artists	17,000	18,000	20,000	−2,000	0,0169
18 Athletes, sportsmen and related workers	10,000	12,000	16,000	−4,000	0,0480
19 Professional, technical and related workers n.e.c.	108,000	114,000	168,000	−54,000	0,1035
20 Legislative officials and government administrators	4,000	5,000	6,000	−1,000	0,0648
21 Managing and higher executive functions exclusive of public administration	163,000	192,000	272,000	−80,000	0,0859
30 Clerical supervisors	17,000	21,000	32,000	−11,000	0,1147
31 Government executive officials	16,000	17,000	23,000	−6,000	0,0542
32 Stenographers, typists and card- and tape-punching	147,000	184,000	149,000	35,000	0,0550
33 Bookkeepers, cashiers and related workers	291,000	315,000	397,000	−82,000	0,0429
34 Computing machine operators	11,000	14,000	8,000	6,000	0,7445
35 Transport and communications supervisors	27,000	29,000	20,000	9,000	0,1664
37 Mail distribution clerks	44,000	47,000	49,000	−2,000	0,0022
38 Telephone and telegraph operators	18,000	17,000	12,000	5,000	0,1524
39 Clerical and related workers n.e.c.	379,000	454,000	456,000	−2,000	0,0000
40 Managers (wholesale)	30,000	35,000	42,000	−7,000	0,0274
41 Managers (retail trade)	17,000	19,000	22,000	−3,000	0,0167
42 Working proprietors (wholesale)	22,000	26,000	30,000	−4,000	0,0153
43 Shopkeepers, street vendors	82,000	93,000	39,000	54,000	1,9630
45 Sales supervisors, and buyers	45,000	49,000	53,000	−4,000	0,0053
46 Technical salesmen, commercial travellers and manfacturers' agents	50,000	53,000	50,000	3,000	0,0049
47 Insurance, real estate, securities and business services	35,000	42,000	56,000	−14,000	0,0589
48 Salesmen, shop assistants and related workers	232,000	290,000	440,000	−150,000	0,1164
50 Managers (catering and lodging services)	13,000	16,000	28,000	−12,000	0,1624
51 Working proprietors (catering and lodging services	30,000	34,000	20,000	14,000	0,5189
52 Housekeeping and related service supervisors	10,000	10,000	9,000	1,000	0,0112
53 Cooks, waiters, bartenders and related workers	108,000	128,000	183,000	−55,000	0,0885
54 Maids and related housekeeping working n.e.c.	181,000	195,000	254,000	−59,000	0,0543

18. For certain occupations the actual outcome in 1992 is below the lower bound of 5,000 persons used by the CBS. Therefore they are not included in the table. However, as far as the actual outcomes are longer than 2,500 they have been incorperated in this evaluation study. Numbers below 2,500 could not be incorperated, but for the overall picture these small occupations are not important.

Tabel 4.7. Quality of the forecasts of expansion demand per occupational class (continued)

CBS-occupational class	number 1985	forecast 1992	outcome 1992	forecast error	loss
55 Building caretakers, charworkers, cleaners and related workers	137,000	189,000	165,000	24,000	0,0223
56 Launderers, dry-cleaners and pressers	10,000	10,000	8,000	2,000	0,0299
57 Hairdressers, barbers, beauticians related workers	31,000	31,000	41,000	−10,000	0,0620
58 Protective service workers	56,000	61,000	68,000	−7,000	0,0104
59 Service workers n.e.c.	29,000	35,000	36,000	−1,000	0,0028
60 Farm managers and supervisors	6,000	6,000	7,000	−1,000	0,0018
61 Farmers	129,000	134,000	105,000	29,000	0,0818
62 Agricultural and animal husbandry workers	131,000	131,000	156,000	−25,000	0,0271
64 Fishermen, hunters and related workers	4,000	4,000	5,000	−1,000	0,0607
70 Production supervisors and general foremen	101,000	95,000	252,000	−157,000	0,3912
72 Metal processors	13,000	14,000	14,000	0	0,0043
73 Wood preparation workers and paper makers	7,000	5,000	6,000	−1,000	0,0212
74 Chemical processors and related workers	24,000	28,000	29,000	−1,000	0,0009
75 Spinners, weavers, knitters, dyers and related workers	12,000	11,000	9,000	2,000	0,0498
77 Food and beverage processors	60,000	62,000	69,000	−7,000	0,0109
79 Tailors, dressmakers, sewers, upholsterers and related workers	32,000	36,000	36,000	0	0,0000
80 Shoemakers and leather goods makers	8,000	8,000	7,000	1,000	0,0119
81 Cabinetmakers, woodworkers and related workers	18,000	18,000	24,000	−6,000	0,0729
83 Blacksmiths, toolmakers and machine tool operators	41,000	33,000	54,000	−21,000	0,1544
84 Machinery fitters, machine assemblers and precision-instrument makers (except electrical)	155,000	143,000	167,000	−24,000	0,0208
85 Electrical fitters and related electric and electronic workers	110,000	136,000	120,000	16,000	0,0192
87 Plumbers, welders, sheet metal and structural metal prepares and erectors	105,000	112,000	117,000	−5,000	0,0016
89 Glass formers, potters and related workers	9,000	9,000	7,000	2,000	0,1507
90 Rubber and plastics product makers	14,000	14,000	16,000	−2,000	0,0285
91 Paper and paperboard product makers	6,000	6,000	5,000	1,000	0,0214
92 Printers and related workers	43,000	49,000	56,000	−7,000	0,0154
93 Painters	39,000	39,000	46,000	−7,000	0,0245
94 Production and related workers n.e.c.	13,000	12,000	18,000	−6,000	0,0990
95 Bricklayers, carpenters and other construction workers	169,000	174,000	189,000	−15,000	0,0065
96 Stationary engine and related equipment operators	6,000	4,000	5,000	−1,000	0,0383
97 Material handling and related equipment operators, dockers and freight handlers	160,000	172,000	174,000	−2,000	0,0001
98 Transport equipment operators	148,000	160,000	205,000	−45,000	0,0491
99 Labourers n.e.c.	45,000	46,000	37,000	9,000	0,0712

A number of the occupational class forecasts were radically inaccurate. For the shopkeepers and street vendors (occupational class 43), growth of 14% was forecast, but employment actually fell by 53%. For computing machine operators (occupational class 34), 23% growth was expected but employment actually fell by 34%. The working proprietors in catering and lodging (i.e., hotel) services (occupational class 51) were expected to enjoy growth of 12%, whereas the number of jobs fell 35%. These abrupt changes in employment raise the question of whether they are indeed attributable to real changes in employment levels, or whether there may have been drastic changes in the methods of measuring or classifying employment.

The forecast quality

Table 4.8 contains the data on the overall quality of the forecasts of expansion demand by occupational class. Looking at the absolute size of the errors, it can be seen that the forecast contains 73% of the error in the SAB forecast. This improvement is considerably lower than was the case for the forecasts for economic sectors. A comparison with table 4.6, in which the forecasts for economic sectors were translated in terms of occupational classes with the aid of the fixed coefficient matrix, shows that the forecasts for occupational classes are no better than they would have been had this fixed coefficient method been applied in translating the sector forecasts of the CPB.[19] This comparison shows that, for any given quality of forecasts for economic sectors, about half of the forecasting error in the forecasts of expansion demand per occupational class derives from errors in the forecast of the occupational structure. In fact the average loss (0.0738) would be reduced to 0.0358 if forecasts for occupational classes were made on the basis of the forecasts for economic sectors, but with the aid of the correct occupational structure. However the forecasts for economic sectors are based on more recent information (up to and including 1988), while the forecast of the occupational structure was based on data from 1985 and earlier. Thus, had the information available been of more even quality, this ratio would presumably have been rather lower.

Table 4.8. Overview of the quality of the expansion demand forecasts per occupational class

Method	av. loss forecast	av. loss SAB	score	number occupational classes
Absolute	0.0738	0.1006	0.73	75
Relative	0.0727	0.0767	0.95	75
Trend always	0.0290	0.0617	0.47	11
Trend sometimes	0.0151	0.0217	0.69	7
Trend never	0.1096	0.1020	1.07	57

19. The loss for the SAB forecast as the level of occupational classes should in principle be equal to the loss for the fixed coefficient method with the figures for economic sectors and the structure matrix from 1985, in table 4.6. The small difference arises because the second method used the full matrix, economic sector x occupational class, in which more people are placed in a 'remainder' category because either their occupational class, or the economic sector, was not known.

If we consider the relative forecasts, the difference between the employment forecasts per occupational class and the SAB forecast becomes even smaller. The average loss of the forecasts is then as much as 95% of the loss for the SAB forecast.

Section 2.4 lists possible weak points in the methodology which was used, including the assumption that changes in employment at the level of occupational classes are determined by demand and the fact that trends are frequently not extrapolated and the information which might be available in the most recent observation, in particular, is ignored. In the framework of this evaluation, it is difficult to determine whether the assumptions as regarding the extent to which employment levels are determined by demand are correct. In table 4.8, the qualities of the forecasts are grouped together according to how the trend is incorporated in the model. Because separate equations have been estimated for the agricultural and manufacturing sectors on the one hand and the private and public service sectors on the other hand, the trend term may be incorporated in the forecast of employment for a particular occupation either 'always', 'sometimes', or 'never'.[20] The separate forecasting errors for these three groups have only been calculated in relative terms, so that errors in the employment forecasts themselves cannot influence these figures.

The table shows that the incorporation of a trend in the forecasting methodology is indeed important. The occupational classes for which a trend was always extrapolated have the best score, and the category without any extrapolation of trends has the worst score. The average loss of the category with a trend in some cases is smaller than for the category with a trend in every case. These occupational classes also produce a better prediction on the basis of the Same As Before forecast. It is striking that this category contains mainly occupations in which there is relatively little change. It is the occupational classes for which the trend was not extrapolated which were most variable. Presumably it is because of this variability that the trend term was at that time not significant. Statistical significance and economic relevance apparently do not always go together.

The failure to incorporate the trend variable thus also explains much of the mediocre forecasting results for the employment forecasts per occupational class. The occupational classes for which a trend was incorporated had a forecast quality comparable to that of the forecasts for economic sectors. The occupational classes for which no trend was included in fact had a worse forecast quality than the Same As Before forecast. This may indicate that not adjusting the forecasts in accordance with the most recent observations, so as to include the unexplained effects at least up to and including the base year in the forecasts, is indeed a shortcoming, as was assumed in section 2.4.

---

20. Some occupations are found in only one of these two broad areas of the economy. If this is the area in which a trend is incorporated, they fall under the 'always' group, and if it is the area in which no trend is used they fall under the 'never' group.

On the basis of table 4.8 we can also conclude that the incorporation of a trend term would clearly make it possible to improve the quality of the employment forecasts for occupational classes considerably. However it should be borne in mind that the occupational classes for which a trend was always used are presumably the classes for which the trend was most measurable (i.e., significant). If a trend term were also to be used for the other occupational classes, its predictive value would therefore probably be lower than in the cases in which it has already been incorporated. However the introduction of random coefficient models, as was done in the forecasts of 1993, is expected to make it possible to extrapolate less definite trends in a reliable way. This will mean that the trend extrapolation will be more cautious where the trend was less significant during the historical period. It would therefore not be necessary to make an all or nothing choice between using the trend or ignoring it, as was the case for the forecasts being evaluated here.

An explanation of the forecasting error

Table 4.9 explains the standard deviation in the forecasting errors for the employment forecasts per occupational class. The explanation begins with a scale effect, which shows the extent to which the forecasting error is related to the size of the occupational classes. This scale effect is again less than 1, in accordance with the findings in the previous section. This means that the increase in the forecasting error is less than proportional to the size of the occupational class. Large occupational classes thus have more reliable forecasts, in relative terms, than smaller occupational classes.

The second explanatory variable is the method of trend extrapolation. In line with the previous findings, the two groups of occupational classes in which a trend was always or sometimes extrapolated have significantly better results than the group for which no trend was extrapolated.

Finally, a connection has been made between the risk indicators and the magnitude of the forecasting error. It would be expected that the risk indicators incorporated in the overall characterization of the employment prospects of an occupational class would give an indication of the probability that the actual changes in employment levels may differ from the forecast. As regards the dispersion indicator, a wider dispersion should mean a lower probability of the actual changes in employment levels differing sharply from the forecast. This effect was in fact found in the results of the estimations. As regards the sensitivity indicator, greater sensitivity to the state of the business cycle would be expected to increase the risk of a bad forecast, but the effect which was actually found was the reverse. This is probably because, at the level of economic sectors, the forecasts for the manufacturing sectors were better, and it is the occupations which are most represented in these sectors which suffer from the greatest sensitivity to the state of the business cycle. However neither the effect of the dispersion indicator nor the effect of sensitivity to the state of the business cycle was significant.

Table 4.9. Explanation of the standard deviation of the errors in the forecasts of expansion demand per occupational class

Variable	parameter	t-value
Scale	0.84	11.50
Constant	1.03	1.31
Trend always	−0.57	2.65
Trend sometimes	−0.83	3.40
Dispersion index	−0.01	0.66
Sensitivity index	−0.19	0.48

Table 4.10. Over-estimation of changes in the relative expansion demand per occupational class

	underestimation coefficient	t-value
Total	−0.33	1.18
Trend always	0.31	0.37
Trend sometimes	−0.59	0.10
Trend never	−0.85	2.08

Over-estimation of change

Table 4.10 shows the degree to which over-estimation or under-estimation of change was evident in the employment forecasts per occupational class. Once again it is the relative forecasting errors which are considered, so that there is a correction for the forecasting error as regards the total change in employment levels. The only occupational classes for which changes were significantly over-estimated were those for which the trend was not used. Apparently the predicted changes were not very informative, consisting mainly of noise. For the occupational classes for which no trend was used, this noise presumably arises partly from measurement error but mainly from the unjustified decision not to take account of the difference between the most recent observation and the regression line.

Qualitative characterizations

Thus far the evaluation of the employment forecasts has been based on the forecasts in point terms, as given in the report on the forecasting process (De Grip *et al.*, 1989). As has been noted, students are not in fact presented with these point scores, but rather with qualitative characterizations which are derived from the point forecasts. Five categories were defined as follows:

		EOC	≤ 1%	very low
−1%	<	EOC	≤ 9%	low
9%	<	EOC	≤15%	average
15%	<	EOC	≤25%	high
25%	<	EOC		very high

Where EOC is the expansion demand in an occupational class (De Grip *et al.*, 1989, p. 10).

Table 4.11 is a matrix of the forecast and actual outcomes in terms of this qualitative characterization, again using the relative forecasts. The results for the absolute forecasts would be completely dominated by the under-estimation (or incorrect measurement) of the total growth in employment. Of the 75 occupational classes for which the actual outcome is known, 13 have the correct qualitative characterization in absolute terms. Because of the under-estimation there are only 13 occupational classes for which the forecast was higher than the actual outcome, while the outcome was higher than the forecast for 49 occupational classes.

Table 4.11. Characterisations of forecast and actual changes in employment level by occupational classes (relative)

Outcome	very low	low	forecast average	high	very high	total
Very low	**9**	**12**	**6**	3	0	30
Low	3	3	**4**	2	1	13
Average	2	0	**2**	0	0	4
High	1	7	**2**	2	0	12
Very high	1	5	**1**	**7**	**2**	16
Total	16	27	15	14	3	75

Due in part to the difference in measurement methods between the Employee Census and the EBB, the relative forecasts give a more mixed picture of the accuracy of the qualitative characterizations. Some 18 of the 75 occupational classes lie on the diagonal of matrix 4.11. That means that 24% of the occupational classes were given the correct characterization. If we add in the occupational classes for which the qualitative characterization in the forecast was adjacent to the characterization of the actual outcome, then this percentage rises to 61% (46 cases). It can clearly be seen in the table that the omission of the trend variable has also had a strong influence on the qualitative characterizations. This shortcoming has meant that the actual outcomes, on average, are in a more extreme category than the forecasts. This is indicated by the matrix elements which are printed in bold type. This *underestimation curve* contains 45 (60%) of the occupational classes.

This means that, while it is true that quite a lot of characterizations were not entirely correct, the error is generally only that no change, or too little change, has been predicted. But in 20 of the 75 cases the characterization was in the wrong direction. In 14 cases the forecast was for low or very low growth for occupational classes which in fact enjoyed higher or very high employment growth.

Appendix I gives an overview of the characterization and the actual outcome for each occupation. The most extreme bad forecast was for Production supervisors and general foremen (occupational class 70) for which the expected growth in employment was characterized as 'very low', whereas in retrospect the correct characterization would have been 'very high'. There were several other classes for which the expected changes in employment levels were characterized as low or very low, but for which the actual growth in employment turned out to be high or very high. These were the lawyers, public notaries and judges (occupational class 12), authors, journalists, and related workers (occupational class 15) sculptors, painters, photographers and related creative artists (occupational class 16), professional, technical and related workers n.e.c. (occupational class 19), legislative officials and government administrators (occupational class 20), government executive officials (occupational class 31), book-keepers, cashiers and related workers (occupational class 33), transport equipment operators (occupational class 98), fishermen and hunters and related workers (occupational class 64), cabinetmakers and other woodworkers (occupational class 81) and production and related workers not classified elsewhere (occupational class 94). Relatively many of these occupations are strongly represented in the public services sector and the government, areas in which the sectoral forecasts of the CPB also proved to be too low. It is also true that several of these occupational classes are characterized by having many smaller jobs.

For the stenographers, typists and data-typists (occupational class 32), computing machine operators (occupational class 34), service workers not classified elsewhere (occupational class 59) and electrical fitters and related electronics workers (occupational class 85), on the other hand, high or very high growth in employment was expected but the actual growth proved to be low or very low.

Of these occupational classes for which the characterization was in the wrong direction, only one (computing machine operators etc., occupational class 34) was also in the list of the occupational classes with the largest forecasting errors on the basis of the point estimates. The other occupational classes with the largest forecasting errors were all cases in which the actual changes in employment levels had been under-estimated. The growth in employment in these occupational classes was characterized as average, while there finally proved to be either very low or very high growth. This is one more illustration of the fact that omitting the trend term has caused by far the largest part of the problems in the forecasts.

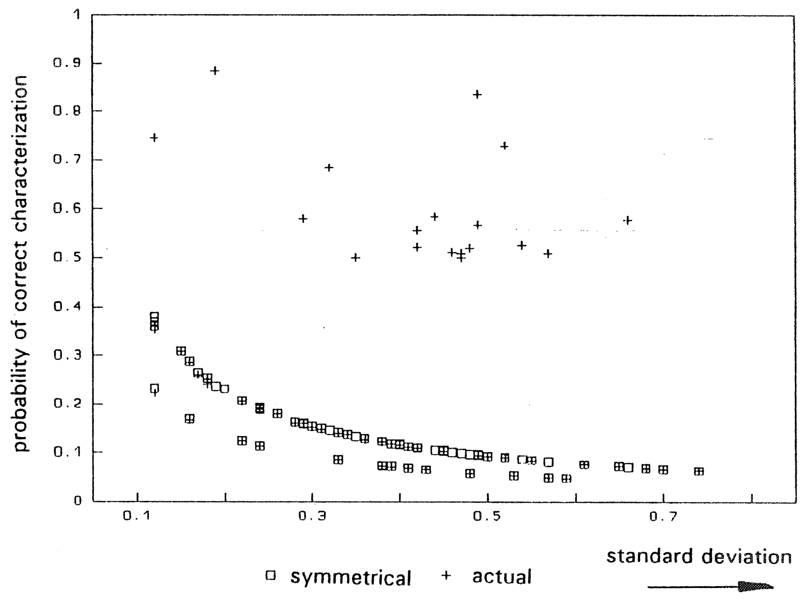
Because of this large forecasting error as a result of ignoring the trend, the observed standard deviation in the forecasting error is very large. Appendix I contains these figures, calculated on the basis of the estimations in table 4.9. On the basis of this standard deviation, it is possible to calculate the probability of the actual outcome having the same characterization as the forecast. As has been said in section 2.10, the characterization method which was used means that this probability is generally smaller than it would be if the intervals were symmetrically placed on either side of the forecast. But the probability of a correct characterization is much higher for the occupational classes in the characterization intervals 'very low' or 'very high', because these intervals have no boundary on one side. The fifth column of the table in appendix I shows what the probability of a correct characterization would have been if the intervals had fallen symmetrically on either side of the point forecasts, and the interval length was the same as that of the intervals which were actually used. The probability that the actual outcome and the forecast would have the same characterization is not very high, varying between 0.05 and 0.36 with a weighted average of 19% (see table 4.12). However the boundaries of the intervals which were actually used are fixed, and the interval does not necessarily fall symmetrically on either side of the prediction. In principle this should lead to a lower probability of a correct characterization. The sixth column of the table in appendix I therefore shows what the probability of a correct characterization is if the forecast is randomly placed in the relevant interval, taking the forecasting error as given. The average probability that a forecast and outcome will fall in the same interval does indeed decline by just over 1%, to 18%. The probability of a correct characterization for many occupational classes is 0, meaning that the forecasting error for that occupational class is so large that the forecast and actual outcome cannot possible fall in one interval using the random placement approach. However the calculation in this case again used intervals at either extreme of the scale whose length was equal to the length of the adjacent intervals. However the extreme intervals in fact had no boundaries on one side, so that the actual probability of a correct characterization in these cases is much higher (seventh column of the table in appendix I). The average probability of an accurate characterization is now 25%, which is reasonably close to the actual result of 24% correct characterizations. The difference is due purely to chance.

Table 4.12. Average probabilities of an accurate forecast using various sorts of intervals to characterize the expansion demand by occupational classes

Type of interval	average probability %
Symmetrical	19
Random position in interval with borders	18
Intervals which were actually used	25
Actual result	24

Figure 4.3 shows the relationship between the standard deviation of a forecast and the probability of a correct characterization, using both symmetrical intervals and the intervals which were actually used at that time. The figure again illustrates the difference between the probabilities of accuracy with symmetrical intervals and with the fixed intervals used for the characterizations in the information system. For the symmetrical intervals, the difference between the 'average' interval (length 6%) and the others (length 10%) is striking. In principle the symmetrical intervals have a higher probability of correct characterizations, but the difference between symmetrical and non-symmetrical intervals is in fact quite small, because of the very large standard deviation of the forecasting errors. The distribution is so flat that it makes little difference whether the interval is precisely symmetrical. However the characterization which was actually used proves to be much better in the case of the extreme characterizations ('very high' and 'very low') because these intervals have no boundaries on one side.

Figure 4.3. Probability of a correct qualitative characterization using symmetrical intervals and the intervals which were actually used, for expansion demand per occupational class



Conclusions

In brief, we can say that the largest shortcoming of the employment forecasts per occupational class was that trends indicating shifts in the occupational structures have in many cases not been extrapolated. The informative value of the most recent observation should also not be underestimated. By choosing a model specification in which the last observation serves as the baseline for the predicted changes, the information which is not contained in the model but is already evident in this observation can be taken into account in the forecasts. In the methodology which is now in use, both of these criticisms have already been satisfied.

As a result of these problems, the accuracy of the characterizations was limited. However the majority of the errors, and the largest forecasting errors, relate to changes in the total employment levels which were ignored in the forecasts. Forecasts and actual outcomes with opposite signs were not so common. The low accuracy points to the possibility that the intervals in the scale which was used were too small. It is true that improved forecasting methods would increase the accuracy, but if we suppose that it is possible to remove about half of the forecasting error (which would be comparable with the forecast quality for the occupational classes for which a trend was always used), the probability of an accurate characterization would still be no higher than 45%. In order to show this uncertainty, it would probably be better to give the employment forecasts for occupational classes in terms of a double characterization, such as 'high to very high', or 'average to low'. The probability of accuracy could then be higher than 70%, while the symmetry of the intervals could to some extent be retained.

The CBS occupational classification system has now been replaced for forecasting purposes by the ROA classification. Because this classification is more suitable for studying the match between education and the labour market and because the forecasts are now compiled in stages (first by occupational segments, and then by occupational classes) the forecast results have probably improved, although it is not possible to assess this in the framework of this evaluation.

4.4. Replacement demand per occupation

The second component which, together with the expansion demand, determines the job openings for each occupational class is the replacement demand. As noted in section 2.5, in the forecasts which were compiled at that time it was in fact not the replacement demand, but rather the flow of people leaving the occupational class which was predicted. All jobs which fell vacant due to the departure of a worker to another occupational class or out of the active workforce were included in these outward flows, so that departing workers who would not be replaced were also included. Jobs which are to be left vacant are not counted in the replacement demand.

To evaluate the forecasts, the actual outflow or replacement demand for the forecast period must first be determined. But this entails the same measurement problems as were encountered in compiling the forecasts. Therefore this evaluation, like the forecasts themselves, is based on flow figures determined by the net method. Simultaneous flows into and out of one defined group are therefore not registered. The many small numbers also raise the possibility of considerable measurement errors in the figures. This means that the errors arising as a result of using the net method cannot be tested and that the evaluation results will be influenced by the low quality of the measurements of actual outflow. Because of changes in the EBB survey, it also proved impossible to obtain outflow figures on the basis of the 1 hour threshold. Therefore in this section outflow figures based on the 12 hour threshold are compared with forecasts for the 1 hour threshold. The most recent forecasts which have been compiled in the framework of the information system (Dekker *et al.*, 1993) have shown that the observed replacement demand using the 12 hour threshold is generally higher than the replacement demand using the threshold which applied at that time. This is apparently because many of the small jobs are held by young people. It is true that the gross outflow for these jobs is high, but it is not observed because in these age groups the flows in are greater than the outflows. Because this effect is not of equal magnitude for all occupational classes, this will distort the evaluation results to some extent.

Table 4.13 gives an overview of the forecasts per occupational class, the actual outcomes, the forecasting error and the loss. As was said in section 3.3, the reference forecast in this case is not the Same As Before forecast, but rather the average predicted outflow. This reference forecast is also shown in the table. In general the predicted outflow is lower than the actual outflow. It would appear that the job-to-job mobility has increased, or that the percentage of those already working who leave the labour market has increased. The largest forecasting errors also relate to this under-estimation of the outflow. The largest aberrations are the forecasts for shopkeepers and street vendors (occupational class 43), housekeeping and related services supervisors (occupational class 52), glass and ceramic workers (occupational class 89), computing machine operators (occupational class 34), farm managers and supervisors (occupational class 60), and forestry workers (occupational class 63). A number of these incorrectly high forecasts also appeared in the evaluation of the changes in employment levels per occupational class (section 4.3) as incorrectly low forecasts. This confirms the suspicion that there is a negative correlation between errors in the forecasts of changes in employment levels and the replacement demand. This negative correlation will in fact mean that these separate components will to some extent cancel each other out in the forecasts of job openings.

Table 4.13. Errors in the forecasts of replacement demand per occupational class

Occupational class	reference	forecast	outcome	forecast error	loss
01 Physical scientists and related technicians	6,000	5,400	11,300	−5,900	0.0232
02 Architects, engineers and related workers	31,600	21,400	71,900	−50,500	0.0444
04 Aircraft and ships' officers	3,000	2,700	8,000	−5,300	0.1419
05 Life scientists and related technicians	5,200	2,600	8,800	−6,200	0.0280
06 Medical, dental, veterinary and related workers	47,300	26,700	37,800	−11,100	0.0008
08 Statisticians, mathematicians and related technicians	9,000	1,500	5,300	−3,800	0.0009
09 Economists	2,700	1,800	5,600	−3,800	0.0236
11 Accountants	2,200	2,000	3,300	−1,300	0.0020
12 Jurists	3,300	1,500	4,300	−2,800	0.0081
13 Teachers	48,800	30,300	26,900	3,400	0.0001
14 Workers in religion	1,800	1,500	3,400	−1,900	0.0302
15 Authors, journalists and related workers	3,500	1,200	4,400	−3,200	0.0067
16 Sculptors, painters, photographers and related creative artists	5,200	5,300	6,000	−700	0.0002
17 Composers and performing artists	3,000	1,500	7,900	−6,400	0.0983
18 Athletes, sportsmen and related workers	1,800	1,600	4,700	−3,100	0.0396
19 Professional, technical and related workers n.e.c.	19,200	9,700	17,500	−7,800	0.0021
20 Legislative officials and government administrators	800	1,200	2,700	−1,500	0.0633
21 Managing and higher executive functions exclusive of public administration	28,900	19,600	47,400	−27,800	0.0104
30 Clerical supervisors	3,100	4,200	14,600	−10,400	0.1044
31 Government executive officials	2,900	3,600	5,800	−2,200	0.0095
32 Stenographers, typists and card- and tape-punching machine operators	26,100	32,400	44,400	−12,000	0.0065
33 Bookkeepers, cashiers and related workers	51,700	46,600	60,400	−13,800	0.0012
34 Computing machine operators	2,000	2,500	7,400	−4,900	0.4270
35 Transport and communications supervisors	4,800	4,600	14,900	−10,300	0.2558
36 Transport conductors	400	600	1,300	−700	0.0283
37 Mail distribution clerks	7,800	4,800	14,500	−9,700	0.0392
38 Telephone and telegraph operators	3,200	3,600	9,200	−5,600	0.2217
39 Clerical and related workers n.e.c.	67,200	60,600	82,100	−21,500	0.0022
40 Managers (wholesale)	5,400	2,400	7,900	−5,500	0.0173
41 Managers (retail trade)	2,900	1,700	9,500	−7,800	0.1336
42 Working proprietors (wholesale)	3,800	3,000	7,000	−4,000	0.0178
43 Shopkeepers, street vendors	14,500	18,000	62,800	−44,800	1.3380
45 Sales supervisors and buyers	7,900	10,700	15,800	−5,100	0.0091
46 Technical salesmen, commercial travellers and manufacturers' agents	8,900	7,600	18,800	−11,200	0.0503
47 Insurance, real estate, securities and business services, salesmen and auctioneers	6,100	5,500	7,900	−2,400	0.0017
48 Salesmen, shop assistants and related workers	41,100	48,700	63,500	−14,800	0.0011

Table 4.13. Errors in the forecasts of replacement demand per occupational class (continued)

	Occupational class	reference	forecast	outcome	forecast error	loss
50	Managers (catering and lodging services)	2,300	4,200	4,400	−200	0.0001
51	Working proprietors (catering and lodging services)	5,300	6,900	17,100	−10,200	0.2708
52	Housekeeping and related service supervisors	1,800	1,800	8,800	−7,000	0.6679
53	Cooks, waiters, bartenders and related workers	19,200	16,200	20,400	−4,200	0.0005
54	Maids and related housekeeping workers n.e.c.	32,000	28,900	29,300	−400	0.0000
55	Building caretakers, charworkers, cleaners and related workers	24,300	17,800	19,700	−1,900	0.0001
56	Launderers, dry-cleaners and pressers	1,800	2,800	6,400	−3,600	0.1998
57	Hairdressers, barbers, beauticians related workers	5,500	4,900	12,300	−7,400	0.0330
58	Protective service workers	9,900	6,100	11,700	−5,600	0.0068
59	Service workers n.e.c.	5,100	5,500	9,000	−3,500	0.0093
60	Farm managers and supervisors	1,100	1,000	5,000	−4,000	0.3778
61	Farmers	22,900	22,000	49,100	−27,100	0.0672
62	Agricultural and animal husbandry workers	23,200	31,400	38,200	−6,800	0.0019
63	Forestry workers	700	1,000	2,500	−1,500	0.3632
64	Fishermen, hunters and related workers	600	800	1,600	−800	0.0249
70	Production supervisors and general foremen	17,800	20,100	40,700	−20,600	0.0067
71	Miners, quarrymen, well drillers and related workers	300	600	1,000	−400	-
72	Metal processors	2,400	2,400	6,400	−4,000	0.0881
73	Wood preparation workers and paper makers	1,300	2,500	3,900	−1,400	0.0494
74	Chemical processors and related workers	4,300	4,100	6,500	−2,400	0.0069
75	Spinners, weavers, knitters, dyers and related workers	2,000	5,300	6,300	−1,000	0.0110
77	Food and beverage processors	10,700	12,100	18,100	−6,000	0.0074
78	Tobacco prepares and tobacco-product makers	400	100	1,600	−1,500	-
79	Tailors, dressmakers, sewers, upholsterers and related workers	5,800	8,400	10,100	−1,700	0.0022
80	Shoemakers and leather goods makers	1,400	1,700	4,400	−2,700	0.1449
81	Cabinetmakers, woodworkers and related workers	3,200	6,600	6,100	500	0.0005
83	Blacksmiths, toolmakers and machine tool operators	7,300	13,600	8,700	4,900	0.0083
84	machinery fitters, machine assemblers and precision-instrument makers (except electrical)	27,600	32,600	38,600	−6,000	0.0013
85	Electrical fitters and related electric and electronics workers	19,500	18,700	38,300	−19,600	0.0267
86	Broadcasting station and sound equipment operators and cinema projections	400	0	2,100	−2,100	-
87	Plumbers, welders, sheet metal and structural metal	18,500	34,500	21,800	12,700	0.0118
88	Jewellery and precious metal workers	400	300	1,900	−1,600	-
89	Glass formers, potters and related workers	1,700	2,400	7,000	−4,600	0.4924
90	Rubber and plastics product makers	2,400	2,800	7,000	−4,200	0.0675
91	Paper and paperboard product makers	1,000	700	3,200	−2,500	0.2182
92	Printers and related workers	7,700	7,800	12,700	−4,900	0.0076
93	Painters	7,000	9,400	7,300	2,100	0.0022
94	Production and related workers n.e.c.	2,300	4,500	4,100	400	0.0006
95	Bricklayers, carpenters and other construction workers	30,000	57,500	30,700	26,800	0.0200
96	Stationary engine and related equipment operators	1,100	1,600	3,100	−1,500	0.1063

Table 4.13. Errors in the forecasts of replacement demand per occupational class (continued)

Occupational class	reference	forecast	outcome	forecast error	loss
97 Material handling and related equipment operators, dockers and freight handlers	28,300	27,100	47,600	−20,500	0.0137
98 Transport equipment operators	26,200	32,500	21,600	10,900	0.0028
99 Labourers n.e.c.	8,000	18,100	33,700	−15,600	0.1819

Table 4.14 gives an overview of the quality of the forecasts of outflows. The predictions are compared with both the actual replacement demand and the actual outflow figures. The fact that the quality score is much better when the forecasts are compared with the outflow, rather than with the replacement demand, indicates that the forecasts being evaluated here do indeed relate to outflows and are therefore not really replacement demand forecasts. It is striking that the average loss of the forecasts, as compared to the replacement demand, is considerably smaller than for a comparison with the outflow. The reference forecast also has a much smaller average loss in this case. This shows that the replacement demand is a much more stable quantity than the outflow. The approach of Willems and De Grip (1990), to extrapolate the replacement demand rather than the outflow, is thus sound.

Table 4.14. Overview of the quality of forecast of outflows per occupational class

Method	average loss forecast	average loss reference	score
Replacement demand			
absolute	0.0065	0.0037	1.78
relative	0.0074	0.0034	2.17
Outflow			
absolute	0.0244	0.0269	0.91
relative	0.0180	0.0181	1.00
Per occupational group (relative)			
0/1 Academics, professionals, and artists	0.0078	0.0102	0.76
2 Executives and senior management	0.0043	0.0015	2.83
3 Administrators	0.0085	0.0106	0.80
4 Commercial personnel	0.0492	0.0604	0.81
5 Service personnel	0.0113	0.0178	0.63
6 Agricultural occupations and fishermen	0.0185	0.0159	1.17
7-9 Tradesmen and industrial occupations	0.0259	0.0152	1.70

The idea that the replacement demand is more important than expansion demand, in forecasts of job openings, is however not confirmed by the figures. Although replacement demand may be a reasonably large component in absolute terms, relevance for forecasting is determined by variability.

A comparison with table 4.8 shows that the variability of the outflow (i.e., the average loss of the reference forecast) using both absolute and relative measures, is only about 25% as high as the variability of the expansion demand. If we switch from outflow figures to replacement demand, this percentage falls to as little as 4%. Thus using replacement demand rather than outflow forecasts is in itself sufficient to ensure a large improvement in the forecast quality, even if the simple reference forecast were to be used. Moreover, if the replacement demand is used in calculating the job openings, negative expansion demand has no influence on the predicted number of job openings. In that case a negative expansion demand can only lead to forecasting errors if growth has incorrectly been predicted. Errors in the magnitude of the predicted decrease in employment then have no effect on the forecast number of job openings.

Figure 4.4. Relationship between expansion demand and outflow, by occupational classes, between 1985 and 1992

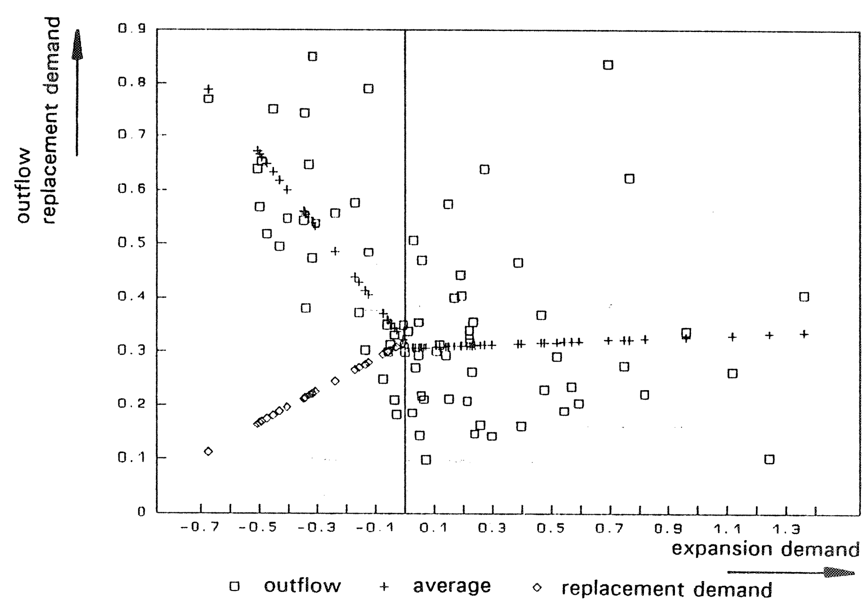


Table 4.15. Correlation between outflow rates and expansion demand between 1985 and 1992, with t-values between brackets

	negative segment	positive segment
Constant	0.32 (2.35)	0.31 (2.01)
Expansion demand	−0.70 (5.05)	0.02 (0.33)

The table also shows that the forecasts of the outflow were of relatively mediocre quality. This would in the first place be because the forecasts are based on an extrapolation of figures for outflows rather than replacement demand. Figure 4.4 shows the relationship between expansion demand and outflow on the one hand, and, on the other hand, expansion demand and replacement demand, for the period 1985 1992.[21] A regression equation has been estimated for both the negative and positive observations of expansion demand. The results can be found in table 4.15, and are illustrated in the figure. It can be seen that the relationship which Willems and De Grip (1990) suggested between outflow and expansion demand is more adequate than the relationship which was assumed in the forecasts of 1989. In the positive segment, no significant relationship was found between expansion demand and outflow, while in the negative segment a negative relationship was indeed found. However the slope coefficient for the regression is larger than the value of -1 which Willems and De Grip assumed. This may be caused by measurement errors in the expansion demand, but it might indicate that a fall in employment levels is in fact partly achieved at the expense of those who are entering the labour market.

Table 4.14 shows a break-down of the forecasting errors per economic sector. The forecasting method which was used scores relatively badly for the management occupations (only two occupational classes), and for the agricultural and manufacturing occupations, which are sensitive to changes in the business cycle. In contrast, good scores were recorded for the commercial, services, and administrative occupations. Relatively many women work in these occupations. Thus, in comparison with the reference forecast, the forecasting method would seem to be good for predicting mainly the more systematic component in the outflow rates, but to be inadequate as regards the effects of the business cycle.

An explanation of the forecasting error

The supposition that the outflow curve has the shape shown in figure 2.2 is implicit in the outflow forecasts. If, as has been shown in the analysis above, the actual outflow curve is different, the forecasts of the outflow will exhibit large forecasting errors particularly where the expansion demand in the forecast period differs markedly from the expansion demand in the period covered by the observations. Therefore in table 4.16 the absolute difference between the two growth rates is incorporated as an explanatory variable for the standard deviation in the forecasting error, with a correction for the total growth in employment in the period concerned, because this overall correction has also been made to the forecasts. The table shows that this factor is indeed a very important and very significant explanatory variable for the forecasting errors as regards the outflow per occupational class.

The scale effect is in accordance with findings regarding the expansion demand per occupational class and economic sector. Large occupational classes are forecast relatively better than smaller occupational classes. The risk indicators are also included as explanatory variable of the forecasting

---

21. Because of the different thresholds for the hours worked per week, the expansion demand growth has been reduced by 15% so that the growth figures for the two definitions of employment are comparable.

error, as was done for the expansion demand per occupational class, but in contrast to the analysis for the expansion demand, both risk indicators do have a significant relationship with the forecasting error. Where the sensitivity to the state of the business cycle is higher, the forecast of the outflow was worse, and for occupational classes with greater switching opportunities, the outflow was forecast better. The greater switching opportunities apparently mean that people are able to switch to other economic sectors in the event of employment problems, so that they can continue to work within their own occupation.

Table 4.16. Explanation of the standard deviation in the errors in the forecasts of outflow per occupational class

Variable	parameter	t-value
Scale	0.78	12.26
Constant	−0.63	0.83
Sensitivity	0.08	5.04
Switching	−0.60	2.64
Changes in expansion demand	13.09	6.05

Over-estimation of divergences

The extent of over-estimation or under-estimation of the observed changes was also examined. Because the reference forecast for the replacement demand is not the same as the situation in the past (the SAB), but rather the average predicted replacement demand, it is in this case not the under-estimation or over-estimation of changes which is analyzed, but rather the under-estimation or over-estimation of divergence from the average. Table 4.17 shows that these divergences were over-estimated in the forecasts of the flows out of occupational classes. As noted in section 3.4, this phenomenon should not be interpreted as if the actual changes were contained in the separate forecasts, but with a multiplier. The negative sign of the under-estimation parameter means that the forecast consists of an one part which is an informative forecast and an additional error term. By restricting the predicted changes to 40% of their magnitude, both the informative value of the forecasts and the noise are reduced. However at first — up to the 40% point — the reduction in the noise dominates, so the net effect is an improvement.

Table 4.17. Over-estimation of changes in the forecasts of relative outflows per occupational class

	under-estimation coefficient	t-value
Total	−0.60	5.10

This noise in the forecasts might equally well be caused by the fact that the forecasts are based on observations from only one historical period, or because large changes may have occurred in the employment for the occupational class. This would mean that the average outflow forecast is (under average conditions) more valuable than outflow forecasts for particular occupations which are based on movements in employment levels which have since changed.

Qualitative characterizations

As with the expansion demand, a qualitative characterization was assigned to the replacement demand per occupational class, so as to make it possible to interpret the forecast better and to show the uncertainty which is contained of the forecast. The classification of the qualitative characterizations was as follows:

	ROC	≤	10%	very low
10% <	ROC	≤	17%	low
17% <	ROC	≤	19%	average
19% <	ROC	≤	25%	high
25% <	ROC			very high

Where ROC stands for replacement demand (thus, in fact, the outflow) per occupational class (De Grip *et al.*, 1989, p. 10).

Table 4.18. Characterizations of the forecast and actual outflows per occupational class

			forecast			
Outcome	very low	low	average	high	very high	total
Very low	2	2	0	1	0	5
Low	4	7	1	2	3	17
Average	0	2	1	2	0	5
High	1	9	1	4	4	19
Very high	4	9	1	10	9	33
Total	11	29	4	19	16	79

Table 4.18 shows how particular characterizations have worked out in actuality. 23 of the 79 occupational classes (29%) were assigned a correct characterization of the outflow. 26 occupational classes were given a predicted qualitative characterization which was one interval too high or too low. Thus in all 62% of the occupational classes were assigned a completely correct or almost correct characterization. However the number of occupational classes that were assigned a characterization which pointed in quite the wrong direction is strikingly high: 29 occupational classes in all (37%). In 6 cases a high or very high characterization was forecast, where it should have been low or very low. In 23 cases the reverse mistake occurred. The four most extreme characterization errors were for composers and performing artists (occupational class 17), managers in the retail trade (occupational class 41), tobacco preparers and tobacco product makers (occupational class 78) and broadcasting station and sound equipment operators and cinema

projectionists (occupational class 86). These mistakes would appear to be caused principally by uncertainties in the measurements. The latter two occupational classes, in particular, are very small. This shows once again that the stability of the forecasts requires more attention.

Appendix II lists the predicted and actual characterization of the outflow for each of the occupational classes, along with the standard deviation per occupational class, the probability of a correct prediction if symmetrical intervals are used, and if the forecast is randomly placed within this interval, and the probability of a correct characterization using the intervals which were actually chosen.

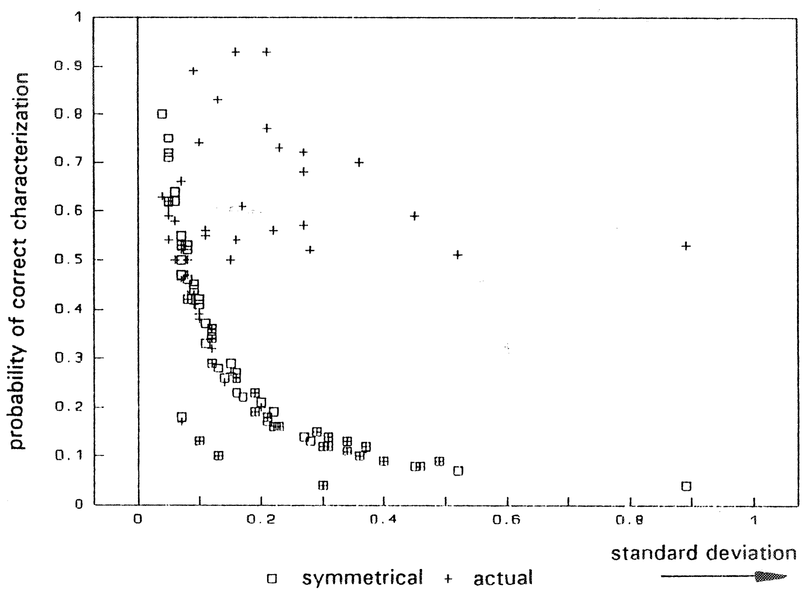
Table 4.19. Average probabilities of an accurate forecast using various sorts of intervals to characterize the expansion demand per occupational class

Intervalkeuze	age probability of being correct %
Symmetrical	46
Random position in interval with borders	34
Intervals which were actually used	47
Actual result	36

Table 4.19 shows that, had symmetrical intervals been used, some 46% of the characterizations could be expected to have been correct. Because the point estimates are randomly placed in the characterization intervals, the actual probability of a correct characterization was only 34%. This arises because a point estimate sometimes lies on the boundary of the interval, so that a small difference between the actual outcome and forecast may be sufficient to make the characterization incorrect. But because the two extreme actual intervals, 'very high' and 'very low', have no boundary on one side, the probability of a correct characterization is in fact higher, at 47%. However the result which was achieved is only 36%. This unexpectedly low score must be the result of chance factors. The 36% of characterizations in the table which were correct is in fact higher than the 29% which is given in table 4.18. This is because the total results in table 4.19 are weighted for the size of the occupational class, and the predictions for the larger occupational classes were better than for the smaller.

Figure 4.5 gives an idea of the relation between the probability of a correct characterization and the standard deviation. The figure gives the probability of a correct characterization using both symmetrical intervals and the intervals which were actually used. Because the standard deviation is substantially smaller than for the expansion demand, it can clearly be seen that a symmetrical interval would yield much higher probabilities of correctness than the interval structure which was used, except in the case of one of the extreme characterizations, 'very low' or 'very high'. The probability of accuracy for occupational classes in these categories is much higher than for the other occupational classes, because the intervals of these extreme characterizations have no border on one side. The occupational classes in each of the other characterization intervals lie approximately on one curve. The highest curves are for the characterizations 'low' and 'high', for which the interval widths are 6 and 7% wide, respectively. The 'average' characterizations, in contrast, have an interval with a width of only 2%. The probability of a correct characterization is clearly considerably lower in this case.

Figure 4.5. Probability of a correct qualitative characterization using symmetrical intervals and the intervals which were actually used, for replacement demand per occupational class



It is also striking that the standard deviations of the outflow forecasts are much more widely distributed than for the forecasts of expansion demand. This is because the occupational classes for which the expansion demand has not changed markedly are much better forecast than the other occupational classes. If the forecasting method can be modified to overcome this defect, and make accurate allowance for the correlation between changes in employment levels and outflow, the probability of accuracy would clearly go up to an acceptable level. If measures were also taken to reduce the uncertainty in the measurements it should be possible to achieve a 70% reliability rating with an interval length of 6-7%.

Conclusions

The analysis in this section shows that the replacement demand forecasts which are being evaluated here had two important shortcomings. The first was that insufficient consideration was given, in extrapolating the outflow coefficients, to the relation between the outflow and changes in employment levels. The current methodology, in which the extrapolation is based on replacement demand, is already a considerable improvement in this respect. Further studies of the outflow curve should enable this method of determining the replacement demand to be improved even further.

The second shortcoming was that the forecasts were unstable because they were based on outflow observations from only one period. Moreover, these observations were in themselves unstable because they were based on rather small fractions of the sample. By basing the forecasts on observations spread over time — which should be possible now that results are available from the Employee Census and EBB survey over several years — and by taking explicit account in the methodology of measurement errors, it may also be possible to improve the forecasts in this respect.

4.5. Job openings per occupation

The opportunities for new entrants to the labour market to find work within a particular occupational class are determined by the job openings, i.e., the sum of the expansion and replacement demand. But because the relation between these two components of demand and the total number of job openings is a matter of definition, there is no point, so far as improving the methodology goes, in making any detailed evaluation of the forecasts of the job openings themselves. The evaluation in this section can therefore be limited to determining the total error in the forecasts of job openings per occupational class, along with an appraisal of the probability that the characterizations will be correct. Because there is no separate specification of the standard error in the predictions of job openings, and also because any qualitative characterization of job openings would in fact be determined by the characterizations of the separate components, there will also be no need to analyze these characterizations thoroughly.

The quality of the forecasts of the number of job openings is naturally determined by the quality of the separate components. However, it was suggested in section 2.5, that there may be a negative correlation between the expansion demand and the flows of workers out of the occupational class. If employment is falling, the outflow should increase. That also means that, if the expansion demand falls more rapidly than was predicted, the outflow will be greater than the forecast. Therefore a negative correlation would also be anticipated between the forecasting errors for the two components of demand. Large forecasting errors for one component are then cancelled out by small errors in the other component of the forecast of the total demand for new entrants to the labour market. Table 4.20 shows that there is indeed a negative correlation. If the two errors had been mutually independent, the total average loss for the job openings would have been 0.0907, the sum of the average losses of the two components, but in fact there was a negative correlation of -0.0142, so that the average loss in the forecasts of the job openings per occupational class was in fact 0.0623.[22] The average loss in the forecasts of the job openings is thus even smaller than the average loss for the expansion demand alone.

---

22. Because $\mathrm{Var}(a+b) = \mathrm{Var}(a) + \mathrm{Var}(b) + 2\,\mathrm{Cov}(a,b)$.

Table 4.20. Composition of the average loss in the forecasts of job openings per occupational class

Component	average loss
Expansion demand	0.0727
Outflow	0.0180
Correlations	−0.0142
Total job openings	0.0623

The classification of the qualitative characterizations was based on a division into five intervals as follows:

	TJO	≤ 14%	very low
14% <	TJO	≤ 24%	low
24% <	TJO	≤ 34%	average
34% <	TJO	≤ 44%	high
44% <	TJO		very high

Where TJO stands for the total job openings per occupational class (De Grip *et al.*, 1989, p. 11).

Table 4.21 shows a matrix of the qualitative characterizations which were assigned and the actual outcomes. It can be seen that the probability that the characterizations of the job openings will be accurate is higher than for the expansion demand forecasts. This is due, on the one hand, to the smaller forecasting error for the job openings and, on the other hand, to the fact that the intervals used for assigning the qualitative characterizations were broader. It can be seen from the table that 26 of the 79 occupational classes (33%) were assigned an exactly correct characterization. If we add in the occupational classes for which the actual outcome was in an interval immediately beside the forecast characterization, we can say that a total of 50 of the 79 occupational classes were assigned a reasonably correct characterization (63%).

However there are also a number of very incorrect forecasts. The occupation classes of authors, journalists, and related workers (occupational class 15) and production supervisors and general foremen (occupational class 70) were predicted to have very low numbers of job openings, but the numbers were in fact very high. The opposite extreme case did not arise. A total of 12 occupational classes which were predicted to have relatively low or very low numbers of job openings should in hindsight have received a high or very high characterization. In contrast, there were another 12 occupational classes which were predicted to have a large or very large number of job openings, but for which the number of job openings proved in fact to be low or very low. Thus a total of 30% of the occupational classes were assigned a characterization which in retrospect lay on the wrong side of 'average'.

Table 4.21. Characterisations of the forecast and actual job openings per occupational class (relative)

Outcome	forecast very low	low	average	high	very high	totaal
Very low	4	2	1	2	0	9
Low	2	8	3	6	4	23
Average	2	5	8	3	0	18
High	4	3	3	2	0	12
Very high	2	3	3	5	4	17
Total	14	21	18	18	8	79

To sum up, the average loss in the forecast of job openings per occupational class is relatively low in comparison with the error in the expansion demand forecasts. This is due to the negative correlation between forecasting errors in the expansion and the replacement demand. If the suggested improvements are made in the forecasting approach for both components of the demand, the quality of the job opening forecasts could be further improved.

4.6. Replacement demand per type of education

The forecasts of replacement demand per type of education were based, in the rather simple manner outlined in section 2.7, on the number of working people aged 55 or older. This was because the data on the age structure which was used for the replacement demand per occupational class was not available then in terms of workers' educational backgrounds. A forecast was made of what is now called 'outflow', just as was done for the occupational classes. 'Vacancies' created by the departure of workers who would not be replaced are therefore again counted in under this definition.

The English translations of the names of the types of education are the same as those which were used in the English translation of the 1989 forecasts, with minor grammatical corrections and some abbreviations, except that SOI 361 was then called 'Lower business education, secondary school for tradesmen (lower level), practical training for clerks and salesmen at the school for domestic science and technique', SOI 381 was then called 'School for domestic science and technique exclusive of training for clerks and salesmen', SOI 481 was 'External care, services as taught at the secondary school for medical and other services or the INTAS', and SOI 506 was 'Training college for primary and pre-primary school teachers, secondary-school teacher training, new style'.

Table 4.22. Errors in the forecasts of replacement demand per occupational class

	Education	reference	forecast	outcome	forecast error	loss
301	General secondary education	58,300	65,200	86,800	−21,600	0.0016
321	Junior agricultural education	16,600	19,600	29,700	−10,100	0.0118
331	Junior technical education	78,800	88,100	114,100	−26,000	0.0025
341	Junior transport, communications and traffic education	5,200	4,200	9,000	−4,800	0.0132
351	Junior medical and paramedical education	500	300	2,200	−1,900	0.2759
361	Lower commercial and administrative educ. (3 types)	22,800	26,900	52,300	−25,400	0.0319
381	School for domestic sciences and tech (exc. 361)	21,400	25,300	37,900	−12,600	0.0024
391	Business security and surveillance training	900	700	2,000	−1,300	0.0259
401	General secondary education, int. and higher levels	35,900	35,700	58,700	−23,000	0.0038
406	Training for driving instructor, sports coach	5,500	4,400	11,700	−7,300	0.0547
421	Senior agricultural education	14,900	18,600	19,100	−500	0.0000
431	Senior school for laboratory science	1,000	1,100	5,200	−4,100	0.1254
436	Senior technical training	87,200	86,600	65,700	20,900	0.0009
441	Senior education in transport, communications etc.	6,600	9,400	11,600	−2,200	0.0023
451	Training of nurses and medical receptionists (MDGO)	15,700	12,700	14,800	−2,100	0.0003
452	Secondary school for medical lab. science, higher level	3,300	2,900	7,300	−4,400	0.0280
453	Training for medical clerks etc.	600	800	800	0	0.0000
454	Training for ward onderlies etc.	6,100	4,500	5,100	−600	0.0001
461	Senior retail school and intermediate business educ.	82,800	87,400	106,900	−19,500	0.0008
466	Inter. business educ., dept. of management studies	6,900	7,300	10,900	−3,600	0.0054
471	Social work and welfare work (MDGO)	4,300	3,500	5,900	−2,400	0.0021
481	Intermediate health care educ., MDGO/INTAS	22,700	22,500	22,100	400	0.0000
483	Secondary hotel and catering school, hairdresser's sch.	5,200	6,800	8,100	−1,300	0.0007
486	Fashion design etc.	300	300	2,100	−1,800	0.1622
506	Primary, pre-primary, and secondary teachers training	36,600	27,300	28,700	−1,400	0.0000
511	Training for interpreters and translators	800	600	2,500	−1,900	0.0977
516	Training for pastoral work etc.	600	600	2,200	−1,600	0.1983
521	Agricultural college	1,700	1,400	4,100	−2,700	0.0248

Table 4.22. Errors in the forecasts of replacement demand per occupational class (continued)

	Education	reference	forecast	outcome	forecast error	loss
531	Laboratory college	3,000	2,400	2,700	−300	0.0001
536	Technical college	15,100	12,200	20,500	−8,300	0.0062
541	Transport, communication and traffic college	3,200	4,600	6,100	−1,500	0.0036
551	Nursing college, physiotherapy college	7,500	5,100	6,600	−1,500	0.0003
552	College for medical laboratory science	2,800	1,900	7,800	−5,900	0.0715
554	Dietetics college, etc.	300	200	1,700	−1,500	0.1138
561	Business college, excl. admin. and fiscal studies	15,400	14,400	14,600	−200	0.0000
562	Courses for ergonomists, management science HTS etc.	500	300	800	−500	0.0014
566	Business science college, legal and adm. studies	2,000	2,000	4,700	−2,700	0.0166
571	College of social studies, library studies	10,800	7,400	8,000	−600	0.0000
583	Hotel college	500	600	2,200	−1,600	0.0811
586	Art academy, academy of dramatic art	4,400	3,600	7,000	−3,400	0.0132
606	Teacher training (highest level)	2,300	1,600	6,300	−4,700	0.0575
611	Language and literature (university)	2,700	2,100	4,800	−2,700	0.0040
616	Theology (university)	1,300	1,200	3,800	−2,600	0.1262
621	Agricultural and domestic sciences	800	600	3,500	−2,900	0.1006
631	Mathematics and physics	3,600	2,700	5,400	−2,700	0.0050
636	Technical sciences	6,000	4,100	13,900	−9,800	0.0314
651	Medical sciences	5,800	4,000	8,500	−4,500	0.0080
652	Pharmacy	300	200	900	−700	0.0198
661	Economics and business administration (B.A.)	3,400	2,700	6,700	−4,000	0.0059
662	Econometrics, actuary and management (B.Sc.)	600	400	1,400	−1,000	0.0245
666	Law	4,500	3,100	4,300	−1,200	0.0005
671	Socio-cultural sciences	6,300	4,300	6,000	−1,700	0.0004
686	Fine arts	300	200	2,900	−2,700	0.2532

Table 4.22 shows that this way of dealing with the age structure led to definite under-estimation of the outflow. Almost all forecasts under-estimated the actual outflow. The largest loss was for junior medical and paramedical education (SOI 351), fine arts (SOI 686), training for pastoral work etc. (SOI 516) and fashion drawing (SOI 486). For each of these the actual outflow figures were more than 100% higher than the predicted figures. The fact that many of the students for all four of these types of education are women suggests that the method which was employed is especially inadequate for women's courses, which is not very surprising since this methodology does not take any account of the withdrawal of women from the labour market to care for their own children. It is therefore important that a distinction be made between men and women, and (especially for the women's courses) that the outflows of younger workers should also be accounted for. Both of these improvements have already been incorporated in the current methodology.

Table 4.23. Overview of the quality of forecasts of replacement demand per type of education

Method	average loss forecast	average loss reference	score
Replacement demand			
absolute	0.0021	0.0019	1.09
relative	0.0020	0.0017	1.18
Outflow			
absolute	0.0056	0.0060	0.94
relative	0.0039	0.0040	0.96
Educational level (relative)			
lower	0.0020	0.0032	0.63
intermediate	0.0033	0.0031	1.04
higher	0.0045	0.0057	0.80
university	0.0131	0.0092	1.43

Table 4.23 provides an overview of the average loss in these forecasts, using a constant outflow rate equal to the average predicted outflow as a reference forecast. The scores prove that the forecasts were indeed focused on the outflow. As was the case for the outflow for economic sectors, the average loss is in fact lower if the forecasts are compared with the figures for replacement demand. This implies once again that replacement demand is a much more stable quantity than outflow. In section 2.7. it was assumed that the relationship between expansion demand and outflow would be less strong for types of education than for occupational classes. A comparison with table 4.14 shows that, for types of education, the difference between the forecasting error for the outflow and for the replacement demand is smaller than it was for occupational classes. For types of education, the expansion demand is negative much less often, so that the differences between outflow and replacement demand are smaller.

Table 4.23 also shows how the average loss differs for the various educational levels. From this it can be seen that the average loss increases as the educational level rises. That is also true for the reference forecast, so that the sequence is not so clear-cut when we examine the scores. Apparently the outflow is more variable at the higher educational levels. The relationship between the expansion demand and the outflow for the various types of education, for the period 1985-1992, is illustrated in figure 4.6. As already noted, only a few types of education have suffered a negative growth in employment. For this reason there was no point in estimating separate regression lines for the positive and negative segments. In the positive segment there is a very slight negative correlation, but this is not significant (see table 4.24). This correlation is mainly due to a number of strongly divergent results. Some types of education are fairly new, so that the flows of young workers with these educational backgrounds into the labour market are high, while there is almost no outflow. The correlation which appears apparently arises because the regression is incorrectly based on a correlation between outflow and *ex post* expansion demand, when in fact the correlation with *ex ante* expansion demand should have been estimated.

Figure 4.6. Relationship between expansion demand and outflow, by types of education, between 1985 and 1992

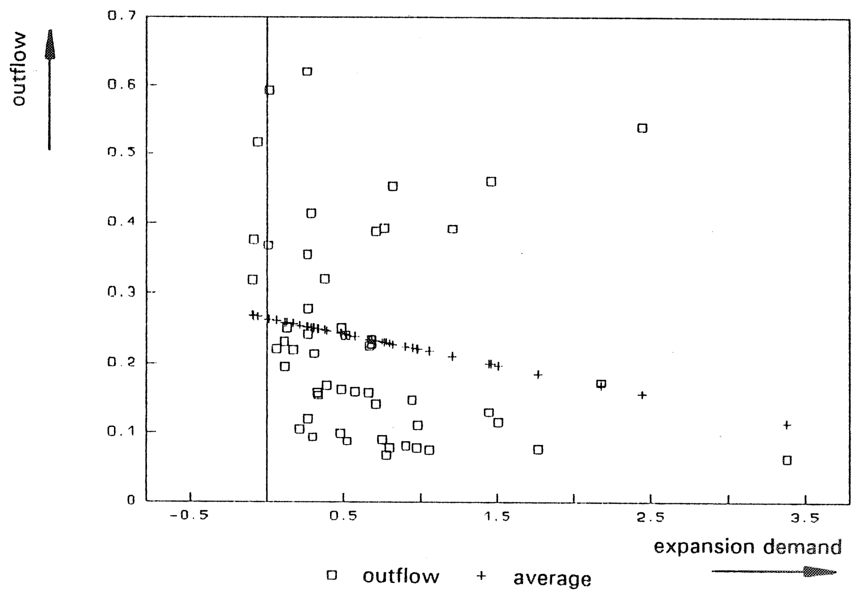


Table 4.24. Correlation between outflow rates and expansion demand, by types of education, between 1985 and 1992

	parameter	t-value
Constant	0.26	1.83
Expansion demand	−0.04	1.47

An explanation of the forecasting error

Table 4.25. Explanation of the standart deviation in the errors in forecasts of outflow per typer of education

Variabele	parameter	t-value
Scale	0.48	7.06
constant	2.80	3.86
Switching occupations	0.45	0.80
Switching sectors	0.18	0.76

Table 4.25 brings together the standard deviation in the forecasting error and a number of explanatory variables. The scale parameter is very low in comparison with the scale parameter for other components of the forecast model. This means that the larger types of education are relatively well forecast. In contrast to the relationship which was assumed to exist as regards the replacement demand per occupational class, the risk indicators do not have any significant influence on the outflow per type of education. This confirms once again the supposition that, for occupational classes, expansion demand effects account for mobility between occupations and thus for fluctuations in the outflow, but that this effect of expansion demand on replacement demand is of no significance for types of education.

The over-estimation of change

Table 4.26. Over-estimation of changes

	over-estimation of changes	t-value
Total	−0.11	0.42

Table 4.26 gives an overview of the estimation result from the test for the degree of under-estimation of divergences. The parameter is negative so that divergences would appear to have been slightly over-estimated. However, in contrast to the over-estimation of divergences in the forecast of replacement demand per occupational class, this effect is not significant. The approach used, of estimating the outflow on the basis of the number of people aged 55 or more, is presumably no better than the method which was used for the replacement demand by occupational classes, but it is less sensitive to measurement errors and so contains less noise which results in over-estimates of changes. However the method used for occupational classes could be improved by incorporating measurement error explicitly in the analysis, so this technique would also be promising for the forecasts of replacement demand per type of education.

Qualitative characterizations

Table 4.27. Characterizations of the forecasts and outcome for replacement demand per type of education

Outcome	very low	low	forecast average	high	very high	total
Very low	0	4	0	0	0	4
Low	0	10	0	2	0	12
Average	0	4	2	1	0	7
High	0	5	2	2	0	9
Very high	1	17	3	0	0	21
Total	1	40	7	5	0	53

The same classification was used for the qualitative characterizations of the replacement demand per type of education as for the characterization of replacement demand by occupational classes. Because the replacement demand by types of education is on average lower than the replacement demand by occupational classes, since job-to-job mobility is not a factor, a large number of characterizations fall in the 'low' category. For 14 types of education (26%) the characterization was exactly correct. For another 11 types of education the characterization which was assigned was one interval on one side or other of the actual outcome. All in all, 47% of all types of education were given a reasonably correct characterization. It is striking that for 23 replacement demand forecasts which were characterized as low or very low, the replacement demand proved in the event to be high or very high. This is 43% of all forecasts. This is mainly due to the large number of low characterizations. Only one very low predicted replacement demand turned out to be very high. This was for pharmacy (SOI 652). And *vice versa*, only two types of education for which the replacement demand was forecast to be high eventually proved to have low replacement demand. It can be seen from the table that the methodology which was used was quite inadequate to predict variations in the outflow.

Appendix IV gives an overview of the characterizations and the actual outcomes per type of education. It can be seen that the larger types of education were relatively well forecast. Table 4.28, which shows the average probability of a correct forecast weighted according to the size of the types of education, shows that this average probability appears to be much greater than the number of exactly correct characterizations in table 4.27. If symmetrical intervals had been used, the average probability of an accurate characterization would have been 54%. Because the qualitative characterizations which were used mean that the forecast is randomly placed in the interval, this probability falls to 44%. Thus not using symmetrical intervals reduces the probability of a correct characterization by 10 percentage points. This percentage is not higher for the intervals which were actually used only because only one type of education fell in an extreme characterization, in this case 'very low',.

Table 4.28. Average probabilities of an accurate forecast using various sorts of intervals to characterize the replacement demand per type of education

Interval choice	average probability of being correct %
Symmetrical	54
Random position in interval with borders	44
Intervals which were actuelly used	44
Actual result	50

Conclusions

In brief, the use of the number of workers aged 55 or more as the forecast of the outflow is not sufficient to explain the differences between types of education. The method which is now used (Willems and De Grip, 1990) is likely to give a better picture of these differences, especially as it accounts also for the temporary withdrawal of women from the labour market. However more explicit account needs to be taken of the measurement errors which appear in these forecasts. Because there are no large differences between the replacement demand for the various types of education, the characterizations proved to be quite good, especially if they are weighted in accordance with the employment level per type of education. Thus it would appear unnecessary to adjust the width of the characterization intervals, considering the methodological improvements which are possible.

4.7. The supply

As noted in section 2.8, it is not possible to evaluate the supply forecasts which are used in the Information System on Education and the Labour Market empirically, because of the methodology which is used and because there is no adequate data on the actual flows from each type of education onto the labour market. Therefore the evaluation of the supply forecasts in this section will be limited to an empirical evaluation of the Skill forecasts of the number of people in each type of education in the regular education system with a qualification. However this is the most important source of data for the supply forecasts.

Table 4.29. Forecast errors in Skill-forecasts of the number of pupils gaining qualifications

	Education	SAB	forecast '85-'92	outcome '85-'92	verschil	loss
1	Lower General Secondary Education	527,800	453,200	440,600	12,600	0.0008
2	Senior General Secondary Education	295,800	268,300	270,900	2,600	0.0001
3	University Preparatory Education	231,700	234,900	229,500	5,400	0.0006
4	Lower Technical School	306,600	231,900	252,700	−20,800	0.0068
5	Individual Technical Education	51,100	43,000	42,100	900	0.0005
6	Lower Nautical Education	4,200	3,300	3,100	200	0.0042
7	Lower Agricultural School	50,400	38,500	39,100	−600	0.0002
8	Individual Agricultural Education	9,800	10,100	10,000	100	0.0001
9	Lower Home Sciences and Manufacturing Education	160,300	119,500	121,700	−2,200	0.0003
10	Individual Home Sciences and Manufacturing Education	34,300	32,300	29,800	2,500	0.0070
11	Lower Retailing Education	18,200	18,500	16,200	2,300	0.0202
12	Lower Commercial and Administrative Education	74,900	63,300	64,700	1,400	0.0005
13	Intermediate Technical School	66,500	67,900	69,000	1,100	0.0003
14	Other forms of Intermediate Technical Education	9,800	7,600	11,100	−3,500	0.0994
15	Intermediate Laboratory Education	7,700	10,000	8,800	1,200	0.0186
16	Intermediate Nautical Education	6,300	5,400	5,300	100	0.0004
17	Intermediate Agricultural Education	35,000	33,400	340,000	−600	0.0003
18	Intermediate Education, Services and Health Care	110,600	155,000	120,400	34,600	0.0826
19	Intermediate Retailing Education	23,100	35,000	35,200	−200	0.0000
20	Other forms of Intermediate Education for Small Businesses	5,600	7,600	7,800	−200	0.0007
21	Intermediate Commercial and Administrative Education	69,300	90,600	84,700	5,900	0.0049
23	Higher Technical School	63,400	41,500	36,300	5,200	0.0205
24	Higher Laboratory Education	7,000	6,900	7,200	−300	0.0017
25	Higher Nautical Education	9,800	9,700	10,300	−600	0.0034
26	Higher Agricultural Education	7,700	11,300	9,600	1,700	0.0314
27	Higher Commercial and Administrative Education	28,000	39,600	42,000	−2,400	0.0033
28	Higher Health Care Education	28,000	33,300	28,700	4,600	0.0257
29	Higher Socio-cultural Education	37,100	36,000	23,200	12,800	0.3044
30	Higher Fine Arts Education	15,400	19,800	16,600	3,200	0.0372
31	Primary School Teachers Training	16,100	15,900	13,500	2,400	0.0316
32	Teachers Training	21,000	19,800	17,100	2,700	0.0249
33	Humanities	18,900	17,600	25,600	−8,000	0.0977
34	Natural Sciences	13,300	15,800	13,700	2,100	0.0235
35	Law	20,300	27,600	24,200	3,400	0.0197
36	Economics	9,100	16,600	14,600	2,000	0.0188
37	Social Sciences	24,500	32,900	29,200	3,700	0.0161
38	Medical Sciences	17,500	15,900	13,800	2,100	0.0232
39	Engineering	17,500	28,100	21,200	6,900	0.1059
40	Agricultural Sciences	4,200	7,100	5,800	1,300	0.0502

Table 4.29 gives an overview of the Skill forecasts and the corresponding actual outcomes.[23] The chosen reference forecast (the SAB forecast) is the outflow in the school year '85 /'86 multiplied by 7 (the number of years covered by the forecast). The table shows that in general the flows coming from the education system were smaller than was predicted in the Skill forecast.[24] The largest error, in relative terms, was for Higher Socio-Cultural Education, for which the actual outflow was 36% less than the forecast. The actual outcomes for Engineering, Other forms of Intermediate Technical Education and Intermediate Education, Services and Health Care were also much less than the forecasts. However 45% more students came from University Education in the Humanities than had been estimated.

Table 4.30. Overview of the quality of Skill forecasts of the numbers of pupils gaining qualifications

Method	av. loss forecasts	av. loss SAB	score
Absolute	0.0142	0.0646	0.22
Relative	0.0122	0.0142	0.86

Table 4.30 presents a total overview of the errors in the Skill outflow forecasts. Like the CPB's forecasts for economic sectors, these cover only part of the period '85 -'92. However data for actual outcomes has been used up to 1989. In absolute figures, the score of the Skill forecasts is very good, but this result is caused mainly by the sharp reduction in the outflow, which meant that the Same As Before forecast produced very bad results. The relative score, which contains a correction for this overall over-estimation, is not so good. The Skill forecasts reduce the forecasting error, in comparison with the SAB, to 86%. Considering that the pupils who would leave the educational system during the forecast period were already in the schools at the time the forecasts were compiled, this result is somewhat unsatisfactory. However, if the average loss for these forecasts is compared with the results for other components of the forecasting system, the forecasting errors must still be regarded as small. Leaving aside the overall reduction in the flows coming from the education system, there has apparently been little change in the flow patterns. The forecasting errors are, in absolute terms, small, so the supply is apparently a reliable component of the labour market prospects, but these forecasts are little better than could be

23. This evaluation is based on the actual outcomes as recorded in the Reference Estimates of 1992 (Ministrie van Onderwijs en Wetenschappen, 1992). However the classification system used in the Reference Estimates is not entirely the same as that used in the Skill forecasts, so that it has been necessary to make a number of adjustments. In the Reference Estimates, Intermediate Retailing Education (MMO) and Other forms of Intermediate Education for Small Businesses are all combined under MMO. The total flows of pupils from MMO for '88 - '92 has therefore been split between these two sub-categories in accordance with the ration between them in '85 -'87. In the same way, figures for HTNO in the Reference Estimates have been split up into Higher Technical School, Higher Laboratory Education, and Intermediate Nautical Education. The figures in the Reference Estimates for INTAS have been added in to Intermediate Education, Services and Health Care (MDGO).

24. However these Skill forecasts covered a period of only 4 years.

achieved with the SAB method.

An explanation of the forecasting error

Table 4.31 presents the results of an explanation of the forecasting error in the Skill forecasts. These explanation contain only a scale parameter and a constant term. The scale parameter is again somewhat less than 1, which implies that the larger types of education were relatively better forecast.

Table 4.31. Explanation of the errors in the Skil forecasts of the numbers of pupils gaining qualifications

Variabele	parameter	t-value
Scale	0.87	5.38
Constant	−0.56	0.34

Over-estimation of changes

The Skill forecasts were also examined to see whether changes were over-estimated or under-estimated. The results are presented in table 4.32. In both absolute and relative terms, the Skill forecasts significantly over-estimated the actual changes in the numbers of qualified school-leavers. This over-estimation of changes is in accordance with the finding that the Skill forecasts suffer a degree of instability. The way in which the most recent observations are included in the model as restricting factors means that the method is very sensitive to accidental factors in these recent observations. This apparently means that the forecasts contain greater fluctuations than would be expected on the basis of the data.

Table 4.32. Over-estimation of change in the Skill forecasts of the numbers of pupils gaining qualifications

	over-estimation coefficient	t-value
Absolute	-0.17	2.72
Relative	-0.22	3.57

Conclusions

To sum up, the errors in the supply forecasts are reasonably small in comparison with other components of the information system. However this is mainly because the changes in the composition of the supply of school-leavers from regular initial education are quite small. The changes are caused mainly by an overall reduction in the flows coming from the education system. Despite the short forecast period and the fact that the predicted outflow in this period consisted of pupils who were already at school when the forecast was compiled, the error in the Skill forecast is not much smaller than in the Same As Before forecast. The Skill forecast also over-estimated changes, which implies that elements which have no predictive value have been incorporated in the forecast.

4.8. The confrontation of demand and supply

As has been noted, the demand for new entrants from each type of education was determined in these forecasts by the number of job openings, i.e., the sum of the expansion and replacement demand, while the supply was determined by the number of school-leavers with each type of education entering the labour market. Neither of the separate components — i.e., forecasts of job openings and of numbers of school-leavers — can be subject to any direct empirical evaluation. As was said in section 4.7, the problem as regards the supply is that there are no figures available for the actual numbers of school-leavers — that is, no figures at the level of the classification system which has been used for types of education, in which any non-regular types of education which may have been completed are also included.

However, as already indicated in section 2.6, there is a more fundamental problem as regards the empirical evaluation of the number of predicted job openings. This is caused by the distinction between the *ex ante* and *ex post* demand per type of education. In contrast to the demand per occupational class, when dealing with types of education it must be borne in mind that the demand which is manifest in practice is largely determined by supply factors. Because the forecasts of the number of job openings are intended precisely to predict the demand before it has adjusted in response to the supply, it is not permissable to compare figures for observed demand with the forecasts. This problem is particularly acute as regards the expansion demand. Theoretically, the tension between the demand and supply of new entrants could also influence the outflow of those already working, but in practical terms the assumption that the outflow is at most determined by the *ex ante* demand, and not by the supply, seems justifiable.

This report thus contains no empirical evaluation of the forecasts of the expansion demand or the number of job openings per type of education. Such an evaluation would only be sensible if, along with the *ex ante* demand forecasts, the expected adjustment processes — which indicate how many new entrants will ultimately find work — were also explicitly modelled. In fact such *ex post* demand forecasts, in combination with the forecasts of supply, would amount to a prediction of the unemployment figures per type of education.

Although it is impossible to evaluate demand and supply forecasts separately for each type of education, demand and supply together will certainly have an influence on the employment per type of education. The demand which is actual expressed in the market will in fact be determined partly by demand factors and partly by the supply. Therefore, before we go on in section 4.9 to evaluate the LMI as an indicator of the discrepancy between demand and supply in the labour market, this section will examine the extent to which demand and supply forecasts enable us to predict the *ex post* job openings. The procedure here, following Freeman (1980), is to estimate the actual employment per type of education as a weighted average of the forecasts of demand and supply. This method, which does not take the mutual interactions between types of education into account, was used by Freeman to make a confrontation between demand and supply per occupational class. By doing so he established that the employment per occupational class is determined mainly by demand factors. This result lends weight to the supposition in the forecast model being evaluated here, that supply factors at the level of occupational classes have only a limited influence on the *ex post* demand.

Table 4.33. Influences of demand and supply on the *ex post* job openings between 1985 and 1992

	parameter	t-value
Constant	24.76	5.29
Demand forecast	0.32	2.58
Supply forecast	0.68	5.57
Score (= $1-R^2$)	0.88	

Table 4.33 contains the results of a similar regression analysis for the *ex post* job openings per type of education on the basis of the forecasts of the information system. These actually observed job openings are defined as the sum of the *ex post* increase in employment and the actual flows of workers out of the occupational class. In the estimation these job openings, expressed as a fraction of total employment, are explained on the basis of the forecasts for the number of job openings and the predicted flows onto the labour market, but with the limiting condition that the sum of the demand and supply influences on the employment level must equal 1. Because the actual changes in total employment levels were greater than was forecast (due in part to the change in the way the question was asked), the regression contains a constant term which corrects for this under-estimation. The estimation results in table 4.33 therefore relate to the relative forecast quality.

The table shows that the supply influence on the employment level for each type of education is larger than the demand influence. Thus it is indeed important to distinguish between the *ex ante* and *ex post* demand per type of education. Both influences are significant, so that employment levels are apparently also not entirely supply-determined. Thus more people with types of education which are *ex ante* in excess supply will not find work, or will not participate in the labour market, than is the case for the types of education which are in short supply.

However the results in table 4.33 must be used with caution: these results are in fact determined not only by the effects of demand and supply, but also by the quality of the demand and supply forecasts. If either demand or supply were badly forecast — i.e., if the forecasts contained unnecessary noise — the influence of the factor concerned on employment will also be found to be lower in the regression. It must also be realised that a supply element is also included in the demand forecasts which are being evaluated here, and this could also have influenced the results in table 4.33.

The estimation results show that the demand and supply forecasts together were a reasonably good basis for forecasting the changes in employment levels. Considering the uncertainty in the separate forecasts and the fact that the relation between changes in employment levels and *ex ante* demand and supply is in practice more complex than is assumed in the regression which is analyzed here, this result could be said to be satisfactory. A study of the actual relationship between the discrepancy between demand and supply and changes in employment levels could lead to a valuable extension of the information system. However this would have to be differentiated for the various types of education, because it would not be expected that the employment level effects would be evidenced in the same way in every case.

4.9. The labour market indicator LMI

As has been shown by the results in the previous section, the function of the forecasts of the *ex ante* demand is not to predict the actual changes in employment levels. The forecasts of the number of job openings serve — in relation to the forecasts of supply — as an indication of the expected tensions in the labour market. These tensions will lead to adjustments which will be expressed in a deteriorating market position for people with a type of education which is in over-supply and an improving situation for people whose educational background is in short supply.

These expected future labour market prospects are expressed in the information system in the LMI, the quotient of the supply and demand for new entrants to the labour market. A high value of the LMI indicates bad prospects, while a low value indicates good prospects. In order to be able to evaluate the forecasts of the future labour market prospects it is thus necessary to consider whether there is in fact any relation between the LMI and the labour market position, or change in the position, of types of education.

Therefore we require indicators for this evaluation which reveal the labour market position of types of education. However the lack of relevant information means that no adequate evaluation is possible. As of now there is insufficient information available regarding the labour market position of types of education in either 1985 or 1992. In the future the school-leaver surveys, the Higher Vocational Education monitor and RUBS, will presumably be able to be used for the purpose, but because these studies of the labour market positions of school-leavers have only recently been established, they cannot be used in the current evaluation.[25] There is at present also no current data on the unemployment per type of education.

25. However the lack of data for University Education imposes another limitation on the possibilities of using the school-leaver surveys for evaluation purposes.

Nevertheless, to get some impression of the predictive value of the labour market indicator, two indicators have been formulated on the basis of the EBB data which can highlight any change in the labour market position, and unemployment statistics have been analyzed in such a way as to get an approximate idea of the unemployment per type of education, so far as this is possible on the basis of the sources of information which are available. The reconstruction is based on current information from the CPB regarding the total magnitude of unemployment, further dissaggregated to the level required using the figures from the 1991 registers of the Regional Employment Services, which have recently been made available. The two indicators which, in addition to unemployment levels, are intended to indicate changes in labour market positions are changes in the level of the jobs in which people with a particular education are found and changes in the dispersion of people with a given type of education across the occupational classes.[26] An unfavourable labour market situation would be expected to result in people working in jobs at a lower average job level, and to a wider employment dispersion, as the situation forces people to switch to other occupational classes. Thus the job level variable is expected to have a negative correlation with the LMI, and the variable for occupational dispersion would be expected to have a positive correlation to the LMI. Unemployment levels will of course correlate positively with the LMI. In fact it would be expected that changes in the job level and changes in the occupational dispersion would exhibit large correlations.

Table 4.34 gives an overview of the relation between the LMI and the three indicators of the labour market situation.[27] It can be seen that the expected relationship between the predicted labour market situation and unemployment, in particular, does exist. Thus the LMI seems to have been a reasonably good forecast of unemployment statistics. But the table shows that unemployment is not the only way in which labour market disequilibriums are manifest. If all of the predicted over-supply had been expressed as unemployment, the associated parameter would have been somewhat higher than 1. However the relationship between unemployment and the LMI is only significant at a confidence level of 85%, while the other two indicators are not significant. The correlation with changes in the levels at which people are employed has the expected sign, but the correlation with the LMI is very weak. The relationship between change in the occupational dispersion and the LMI had the opposite sign to that which was expected. The regression results have been influenced by the strong relationship between the two adjustment indicators. If the indicator of occupational dispersion is used on its own to explain the LMI, the estimated effect does have the expected sign. The largest shortcoming of this analysis seems to be that there are no good indicators available for the labour market position of people with a particular educational background, or for changes in this position. As has already been noted, in the future data from the school-leaver surveys may alter in this situation.

---

26. The change in the educational level of employment has been measured in accordance with Borghans and Heijke (1993); the change in the occupational spread is measured using the Gini-Hirschman indicator.

27. Because the LMI predicts these indicators, it should in fact be included as an explanatory variable in the regression. But that would mean that the three indicators could not be used simultaneously in a simple regression. Moreover, no unemployment data is available for a large number of types of education. These types of education have been left out of the regression. For the types of education for which the LMI was greater than 10, it has been given a value of 10 in the regression.

Table 4.34 also includes a regression equation in which the LMI used at that time is replaced by an indicator in which the total employment for the type of education concerned is included in both the denominator and numerator, as is currently done in calculating the IFL. This change in the definition improves the results considerably. The score improves by 16%, from 0.91 to 0.76, and the parameter of the unemployment is now clearly significant. Apparently adding the total employment to both sides of the tension indicator has indeed had the expected stabilising effect, increasing the predictive value of the indicator.

Table 4.34. Correlations between predicted LMI and unemployment in 1992 and indicators of changes in the labour market positions between 1985 and 1992

Variable	AMI		'ITA'	
	parameter	t-value	parameter	t-value
Constant	1,31	1,46	0,92	12,13
Unemployment '92	0,40	1,49	0,06	3,03
Verandering functieniveau	−4,83	0,56	−0,17	0,24
Verandering spreiding	−0,18	0,04	−0,02	0,06
Score (= $1-R^2$)	0,91		0,76	

Appendix V contains an overview of the LMIs per type of education which were calculated at that time, along with the estimation of the actual outcome of the LMI on the basis of the regression in table 4.34 and the predicted and actual qualitative characterizations. Table 4.35 compares the forecast and the actual outcome of the qualitative characterizations of the labour market prospects per type of education. In reading these tables it has to be realised that not only the forecasts, but also the estimated actual outcomes contain uncertainty. Discrepancies between the forecast and the actual outcome can therefore arise either from forecasting errors or from inaccuracy in determining the labour market prospects which a type of education actually enjoyed.

Table 4.35. Characterizations of the forecasts and estimated outcomes for the LMI, according to table 4.34

	estimated outcome					
ROA forecast	good	reasonable	moderate	bad	very bad	total
Good	3	5	8	0	0	16
Reasonable	0	7	6	0	0	13
Moderate	0	1	2	0	0	3
Bad	0	1	2	1	0	4
Very bad	0	0	5	1	0	6
Total	3	14	23	2	0	42

The qualitative classification of the LMI was at that time based on a division which differed from the other qualitative characterization systems used in the information system:

0	<	LMI	≤ 1.00	good
1.00	<	LMI	≤ 2.00	reasonable
2.00	<	LMI	≤ 4.00	moderate
4.00	<	LMI	≤ 8.00	bad
8.00	<	LMI	or	
0	≥	LMI		very bad

Despite these measurement problems, the results are reasonably good. For 13 of the 42 types of education (30%), the qualitative characterization in the forecast agrees exactly with the actual outcome. This is 30% of all types of education. If we add in the forecast characterizations which lay one interval on one or other side of the characterization of the actual outcome, we can say that 67% (28 of the 42 types of education) were assigned reasonably accurate LMIs. Moreover there are strikingly few types of education for which the forecast points in the opposite direction to the actual outcome. There was one type of education for which bad prospects were forecast, while the actual outcome was reasonable. This was for teacher training at the highest level (SOI 606), which is very heterogeneous and also difficult to distinguish from other types of education. In fact the interim evaluation (De Grip, Heijke and Berendsen, 1991) had already suggested that the forecasts per type of education were better than the forecasts per occupational class. Apparently not all the errors in forecasting the occupational classes are expressed in the predictions for types of education. The relatively easily predictable supply component also has a positive influence on the forecast results for types of education.

5. CONCLUSION

5.1. Introduction

This report has evaluated the first forecasts which were compiled using the Information System on Education and the Labour Market. These forecasts, compiled in 1989, related to the labour market in 1992, differentiated by types of education and occupational classes. Because the forecast period has now run its course, it has been possible to review these forecasts in a more thorough way than in the first evaluation in 1991 (De Grip, Heijke and Berendsen, 1991). In fact there have in the meantime been substantial changes in the structure and methodology of the information system on the basis of the experience acquired in compiling the forecasts, so that the approach used in the forecasts evaluated here is no longer the same as the current methods. In 1989 the information system was still clearly in its infancy. Nevertheless a critical examination of the structure and forecast methods used at that time has proved very informative.

The previous chapters have presented evaluations of the forecasts for the labour market in 1992, by type of education and occupation, looking first at the structure of the information system and the methodology used in the forecasts. The second step was an empirical evaluation of the forecasts which were compiled at that time. In this chapter the most important findings will be briefly summed up, beginning with an overview of the most important results of the empirical evaluation, followed successively by the structure of the information system and the forecasting method which was used. At each stage the findings in this report will be used as a basis to ascertain what improvements have already occurred in the past years and what further development of the information system in the future could increase the forecast quality and the usefulness of the labour market information it provides even further.

5.2. The results of the empirical evaluation

Table 5.1 gives an overview of a number of key figures from the empirical evaluation as regards the separate components of the forecasting model, beginning with the average loss which shows how large the forecasting errors were for each of the separate forecast components. The lowest average loss was for the replacement demand per type of education, and the average loss for the forecasts of replacement demand per occupational class was also quite low. It is apparently possible to compile relatively reliable forecasts of replacement demand. The Skill forecasts of the numbers of students leaving regular education with a qualification also have a reasonably low average loss. However this figures is not really comparable with the other loss figures, because the forecasting errors are not related to numbers of people working, but rather to numbers of students. By comparing the various demand and supply forecasts with each other it can be seen that the forecast for the expansion demand per occupational class has by far the lowest reliability. As might be expected, this demand component is the most difficult element of the information system to forecast. In contrast, the forecasts of the total number of job openings per occupational class — which consist of the sum of the expansion demand and the replacement demand — are more accurate. This is due to the negative correlation between the errors in the forecasts of the expansion demand and replacement demand.

The score — that is the ratio of the average loss of the forecast to the average loss of the 'Same as Before' reference forecast — shows that the loss in the forecasts of expansion demand per occupational class is not exceptional. In comparison with the reference forecast, the prediction of the expansion demand is thus no worse than other components of the forecast model. The score of the point forecasts for most components could be said to be mediocre. The value is generally somewhat lower than 100, which implies that the forecasts are only a little better than the reference forecast. Despite the low average loss, the score for the forecast of the replacement demand per occupational class is no better than the scores for the other components of the information system. This is because the reference forecasts are also relatively good for the replacement demand, whereas for the expansion demand per occupational class, for example, the Same as Before forecast also exhibited a higher average loss. The analysis has shown that the changeover from outflow forecasts to replacement demand forecasts, which has since been made, will result in considerable improvement in the forecasts. There is also a clear reason for the moderate quality of the forecast of expansion demand per occupational class: the lack of a trend in most of these forecasts. The table shows that if the forecasts for economic sectors had been distributed over the occupational groups using the correct structure matrix for 1992, the score for the forecasts for occupational classes would have been 54%. The score for those occupational classes for which a trend variable was incorporated in the forecasts was 47%.

The explanation of the standard deviation of the forecasting error showed that all components exhibit a scale effect. The scale factor is smaller than 1 in every case, which implies that larger occupational classes and types of education have, on average, smaller forecasting errors in percentage terms. This scale effect is largest for the replacement demand by types of education. This part of the table shows that there are variations in the forecasting error within the forecast components. The forecasts for economic sectors are relatively good as regards the agricultural and manufacturing sectors. A clear relationship was established between the replacement demand and the risk indicators in the information system. The greater the sensitivity to the state of the business cycle and the smaller the switching opportunities, the greater the error in forecasting replacement demand. The replacement demand per occupational class also contained a very large forecasting error in cases in which the expansion demand had changed markedly between the period covered by the observations and the forecast period. This shows that the assumption that the flow of workers leaving an occupational class is not dependent on changes in employment levels at that moment was incorrect.

The following section in the table relates to under-estimation and over-estimation of changes. The CPB forecasts of employment per economic sector exhibit a tendency to under-estimate changes. The actual changes are, on average, twice as large as the CPB had forecast. The forecasts of both expansion and replacement demand for occupational classes for which there was no trend term tended to over-estimate changes. These over-estimates were apparently caused by the high noise level in the forecasts. To avoid these problems in the future, careful account will have to be taken of the influence of measurement errors on the predictions. This has already been done in forecasts of the expansion demand per occupational class (see Borghans and Heijke, 1994).

Finally, the last column of the table shows the results as regards the qualitative characterizations, with the percentage of the forecasts which had exactly the correct characterization and also how many forecasts were assigned a reasonably good characterization, i.e., a characterization differing by no more than one interval from the characterization of the actual outcome. It is striking that the labour market indicator (LMI), which is used as the basis for characterizing the labour market prospects per type of education, has especially good results here. This confirms the findings from the interim evaluation, that this tension indicator gives especially good forecast results. However the evaluation results which are summed up in this column are not derived purely from the quality of the forecasts in themselves: the method of assigning characterization has also had an influence. It has been shown that, for some components, the intervals used (especially for the characterization 'moderate') are very small in comparison with the standard deviation of the forecasting errors. Especially as regards the expansion demand per occupational class, which is difficult to forecast, it would be advisable to change to a more general method of characterization.

The empirical evaluation of the forecasts revealed a number of important shortcomings in the forecasting method. Leaving aside problems due to the lack of data, the factors which have had the largest influence on the forecast results were:

- the inadequate incorporation of trends in the forecasts of the expansion demand per occupational class;
- the fact that no account was taken of the relation between expansion demand and outflow;
- the inadequate consideration which was given to measurement errors in the data which was used.

On the first two points, important changes have since been made. The trend is now incorporated in the model for all occupational classes in a justifiable way, and the replacement demand forecasting method has also been much improved. However, considering the sources of data which are now available, it seems desirable to improve the forecasting method for replacement demand even further. Also, for the replacement demand and a number of other components of the forecasting method, there is still not enough allowance made for the effects of inaccuracy in the data. The only component for which there has since been an adjustment in this area is the expansion demand per occupational class. This adjustment appear to work satisfactorily.

One important finding from the empirical evaluation is that it is extremely important to design models which are highly robust. It is quite possible to compile forecasts for education and occupation at a low aggregation level, but any small oversights in the modelling can cancel out much of their potential predictive power. Despite the errors, the forecasts seem to be reasonably good. The expansion demand per occupational class would seem to be an especially difficult element to predict. In accordance with the findings in the previous evaluation report, the results per type of education are better. If, in addition, the shortcomings which were identified above are taken into consideration, it would appear possible to achieve further improvements in the results. However such improvements have in fact already to a large extent been made, with the changes in the structure and methodology of the information system which have been introduced up to now.

Table 5.1. Overview of a number of key figures as regards the empirical evaluation

Component	average loss (relative)	score		explanation of standard deviation		significant over/under estimation of change		characterizations almost correct	exact
CPB economic Sector forecasts	0,0120		0,47	scale agr./manuf.	0,62 −0,22		1,07		
Idem per occupation									
Structure '85	0,0771		1,01						
Structure '92	0,0418		0,54						
Expansion demand Per occup. class	0,0727		0,95	scale	0,84			61%	24%
		trend always	0,47						
		trend never	1,07			trend never	−0,85		
Replacement demand Per occupational class	0,0180		1,00	scale sensit.ind. switch.ind. change exp.d.	0,78 0,08 −0,60 13,09		−0,60	62%	29%
Job openings Per occupational class	0,0623								
Expansion demand Per type of education	no direct evaluation, because of the difference between *ex ante* and *ex post* expansion demand								
Replacement demand Per type of education	0,0039		0,96	schale	0,48	-		47%	26%
Supply skill	0,0122		0,86	schale	0,87		-0,22		
Supply by Type of education	no direct evaluation, because of the lack of data on school-leavers from non-regular education								
Ex post job openings Per type of education			0.88						
Labour market indicator			0.91					67%	30%

Table 5.2. Recommended changes in the structure of the Information System on Education and the Labour Market which have already been effected or which could contribute to further development

Component	changes already made	recommended changes
Expansion demand Per occupational class		
Replacement demand Per occupational class	- forecast of replacement demand rather than outflows of workers	
Job openings Per occupational class	- calculation on the basis of replacement demand + positive expansion demand rather than outflow + expansion demand	
Expansion demand Per type of education	- forecasts no longer contain supply elements - interaction between types of education is included in model	- further studies of competition between types of education
Replacement demand Per type of education	- forecast of replacement demand rather than outflows of workers	
Labour market indicator LMI per type of education	- short-term unemployed counted as competitors for school-leavers	- further research into which groups are relevant as regards competition with school-leavers - in addition to the labour market prospects per type of education, also make a confrontation between the demand and supply by occupational classes - break-down of the labour market indicator according to how labour market tensions are made manifest

Table 5.3. Recommended changes in the forecasting method of the Information System on Education and the Labour Market which have already been introduced or which could contribute to further development

Component	changes already made	recommended changes
Expansion demand Per occupational class	- changes in choice of explanatory variables - additional information in the last observation is included in the prediction - the trend variable is always incorporated - conservative forecasts from 'random coefficient'	- further improvement in choice of explanatory variables - changes in the treatment of trends in the model due to longer time series - improve 'pooling' structure of model
Replacement demand Per occupational class	- replacement demand rather than outflow reduces the sensitivity to expansion demand - correction for changes in rate of participation - forecast is based on average replacement	- study the relation between replacement demand and expansion demand - effect of measurement errors is explicitly demand over several periods included in forecast model - measurement of people re-entering the workforce on the basis of the retrospective question in EBB
Expansion demand Per type of education	- adjustment because observed data relate to *ex post* demand, but forecasts relate to *ex ante* demand	- further development of model by improving the relationship to explanatory variables - study the competitive position of types of education so as to be able to forecast substitution and crowding-out
Replacement demand Per type of education	- outflow is now measured with net method rather than by number of workers older than 55 - correction for changes in rate of participation - replacement demand rather than outflow means the forecast is less sensitive to expansion demand	- study the relation between replacement demand and *ex ante* expansion demand - measurement of people re-entering the workforce on the basis of the retrospective question in EBB
Supply forecasts		- establish historical series of flow figures, so that forecasting method can be made more systematic

Table 5.3. Recommended changes in the forecasting method of the Information System on Education and the Labour Market which have already been introduced or which could contribute to further development (continued)

Component	changes already made	recommended changes
Labour market Indicator per type of education	- including number of workers increases stability and reliability	- more research as to optimal degree of conservatism for indicator
Qualitative	- systematic determination of interval boundaries	- adjust interval width according to forecast quality - match characterizations of components of the information system to each other, give double characterizations for expansion demand per occupational class
Risk indicators	- correction to indicator of occupational dispersion to allow for jobs in which workers' are under-utilized	- more detailed analysis of indicators of skills sensitivity to business cycle and of occupational dispersion

5.3. Recommended changes in the structure of the information system and forecasting method

Both the evaluation of the structure of the information system and the empirical evaluation have shown which components of the structure of the information system and the forecast models which were used function well, and where some adjustment is required in the structure or the methodology. Because it is especially important to obtain an overall view of the components of the information system which need to be modified, the changes for each component are summarised in tables 5.2 and 5.3 . Table 5.2 contains the changes as regards the *structure of the information system*, which are especially focused on making the labour market data more suitable for the primary goal of providing information for educational and vocational guidance. Table 5.3 presents the points on which changes are required in the *forecasting method*. A distinction is made in the tables between changes which are already in place and changes which could be introduced in the course of the further development of the information system. Not every point covered in the tables will be examined in this section, since a detailed review of the separate points can be found in chapters 2 and 4 of this report.

The structure of the information system has remained broadly unchanged. This structure, as illustrated in figure 1.1, appears to function well. The most important adjustments which have already been made affected the replacement demand and the labour market indicator, the basis of the final demand-supply confrontation per type of education. The changeover from forecasts relating to the flow of workers out of an occupational class to a system based on the replacement requirements would seem to have been a clear improvement both for the forecast quality and in terms of making the forecasts more easily interpretable. In contrast to the forecasts for 1992, the influence of the short-term unemployed is now considered in compiling the labour market indicator. To further increase the ease of interpretation of the forecasts, it would be possible to differentiate the rather abstract tension indicators according to how the labour market reacts to over-supply or shortages of people with a particular educational background. For example, this would make it possible in a concrete case not only to characterize the labour market prospects as 'bad', but also to add an estimate of the degree to which these bad prospects will lead people to accept jobs at lower levels or lower wages, or to higher unemployment (or longer search periods). In this connection it would also be worthwhile examining the relation between the labour market prospects of the various types of education. Substitution and crowding out as a result of tensions in the labour market must by definition influence more than one type of education at a time. In fact a start has been made with including crowding-out processes in the calculations for predicting labour market prospects. These first steps should be given more substance through further research as to the competitive position of types of education, for example on the basis of the information from the school-leaver surveys.

Table 5.3 examines the changes in the forecasting method which have already been effected, and further changes which are suggested for the future. Without discussing each of the separate points in this table, there are two aspects which play a role in many of these points. From the evaluation it was seen that making adequate forecasts for the labour market, differentiated by type of education and occupational class and at the aggregation level which is used within the information system, is in principle certainly possible, but that small imperfections in the forecasting method can sometimes lead to a drastic deterioration in the forecast quality. These imperfections are in many cases caused by interference from measurement errors within the forecasts, or they may arise because the forecasting method which is used is not robust enough. Thus, in the further development of the forecasting model, it is in the first place important that the forecasting models are based on stable relationships. Such an improvement has already been made for the replacement demand, by replacing the strongly fluctuating outflow of workers with the much more robust replacement demand as the basis for the extrapolations. Another adjustment, to base the forecasts of expansion demand more heavily on the most recent observations, could offer a simple way of increasing the reliability of the predictions. In the second place, in the further development of the system account must always be taken of the influence of measurement errors, which can have a large influence on the reliability of the forecasts at the low aggregation level which is maintained. By introducing conservative elements into the estimation and forecasting techniques, the measurement errors could be largely eliminated. The results of the 'random coefficients' model which is now applied in calculated the expansion demand per occupational class and the positive effect which has been achieved by stabilising the labour market indicator (IFL) offer two examples of this.

On the basis of the findings from this evaluation, we can conclude that a number of items should be given priority in the further development of the information system. The first of these is the further development of the forecasting methodology for both the expansion demand and the replacement demand. It is also important to add extra dimensions to the labour market indicator. By using information from the school-leaver surveys it would be possible, as has already been noted, to give an indication of the adjustment effects which can be expected to result from particular tensions in the labour market. This extension would markedly improve the interpretability of the forecasts. Only after these components of the forecasting method have been improved would it be sensible to proceed with possible extensions of the information system, such as making the labour market information suitable for other purposes. In this context one might consider extending the forecast period, or of bringing together and comparing demand and supply at the level of occupational classes, which would increase the value of the information system for those formulating labour supply policies. With such adjustments and extensions, the information system would also be more able to provide information which can help those on the demand side of the labour market to get their bearings.

LITERATURE

Beekman, Th.B.J., R.J.P. Dekker, A. de Grip, and J.A.M. Heijke (1989), *An Explanation of the Educational Structure of Occupations,* ROA-W-1989/3E, Maastricht (also published in *Labour* 5-3, 1992 PP. 151-163)

Bierings, H.B.A., J.C.M. Imbens, and C.A. van Bochove (1990), De definitie van de beroepsbevolking, *Supplement sociaal-economische maandstatistieken*, 91-2, pp. 4-12.

Bishop, H.J., and S. Carter (1991), How Accurate are Recent BLS Occupational Projections, *Monthly Labour Review,* October 1991, pp. 37-43.

BLS (1988), *Projections 2000,* Bulletin 2302, Washington D.C.

BLS (1992), *Occupations, Projections and Training Data,* BLS Bulletin 2401, Washington, DC.

Borghans, L. (1991), De arbeidsmarkt en de invloed van voorspellingen, *Economische Statistische Berichten,* 76-3833, pp. 1113-1116.

Borghans, L. (1992), A Histo-Topographic Map of the Dutch University Studies, ROA-W-1992/5E, Maastricht.

Borghans, L. (1993), *Educational Choice and Labour Market Information*, thesis, University of Limburg, Maastricht.

Borghans, L., and J.A.M. Heijke (1993), *Forecasting the Educational Structure of Occupations: A Manpower Requirements Approach with Substitution*, ROA-RM-1993/2E, Maastricht.

Borghans, L., and J.A.M. Heijke (1994), *Een random-coëfficiëntenmodel voor het voorspellen van de beroepenstructuur van bedrijfstakken*, forthcoming.

Bosworth, D.L., G.J. Evans, and R.M. Lindley (1974), Mechanic Manpower Models, in: J.S. Wabe (ed), *Problems in Manpower Forecasting,* Westmead, pp. 61-84.

CBS (1953), *De voorziening in de behoefte aan onderwijzers bij het lager onderwijs in 1952,* Utrecht.

Clements, M.P., and D.H. Hendry (1993), *Towards a Theory of Forecasting*, unpublished paper.

CPB (1993), *Athena; Een bedrijfstakkenmodel voor de Nederlandse economie*, CPB, Den Haag.

De Grip, A., L.F.M. Groot, and J.A.M. Heijke (1987), *Clustering Occupational Classes by Educational Structure*, ROA-W-1987/2E, Maastricht (also published as: Defining Occupational

Groupings by Educational Structure, *Environment and Planning* A, 1991, vol. 23 pp. 59-85)

De Grip, A., and J.A.M. Heijke (1991), Flexibiliteit versus produktiviteit, *Onderzoek van Onderwijs,* vol. 20-1, pp. 6-8.

De Grip, A., J.A.M. Heijke, and H. Berendsen (1991), *Eerste evaluatie informatiesysteem onderwijs-arbeidsmarkt*, ROA-R-1991/1, Maastricht.

De Grip, A., J.A.M. Heijke, J.W. van Dam, R.J.P. Dekker, and M.H. Wieling (1988), *Aanvulling arbeidsmarktmodule I-See project (wetenschappelijk onderwijs)*, ROA-R-1988/2, Maastricht.

De Grip, A., J.A.M. Heijke, and R.J.P. Dekker (1989), *The Labour Market by Education and Occupation to 1992*, ROA-R-1989/8E, Maastricht.

De Grip, A., J.A.M. Heijke, R.J.P. Dekker, Th.B.J. Beekman, and H.M.M. Peeters (1989), *De arbeidsmarktperspectieven van beroepsklassen en opleidingstypen in 1992: rapportage I-See!*, ROA-R-1989/7, Maastricht.

De Grip, A., J.A.M. Heijke, R.J.P. Dekker, and L.F.M. Groot (1987a), *Labour Market Prospects for Occupations and Academic Studies in 1992,* ROA-W-1987/1E, Maastricht.

De Grip, A., J.A.M. Heijke, R.J.P. Dekker, L.F.M. Groot, and L.A. Vos (1987b), *De arbeidsmarktperspectieven van studierichtingen in het wetenschappelijk onderwijs, Arbeidsmarktmodule I-See project*, ROA-R-1987/3, Maastricht.

De Grip, A., J.A.M. Heijke, and L.A. Vos (1987), *Inventariserend onderzoek arbeidsmarktmodule I-See project*, ROA-R-1987/1, Maastricht.

De Grip, A., and E.J.T.A. Willems (1992), *De vervangingsvraag naar beroepsklasse tot 2000*, OSA-W 96, Den Haag.

Dekker, R.J.P., and A. de Grip, *Vergelijking tussen de CBS- en de ROA-beroepenclassificatie*, ROA-W-1992/3, Maastricht.

Dekker, R.J.P., A. de Grip, Th.B.J. Beekman, P.J.E. van de Loo, M.H. Wieling, and E.J.T.A. Willems (1990), *Rapportage I-See! 1990*, ROA-R-1990/6, Maastricht.

Dekker, R.J.P., A. de Grip, H. Berendsen, M.H. Wieling, and E.J.T.A. Willems (1992), *Methodiek en structuur arbeidsmarktmodule I-See! 1991*, ROA-W-1992/1, Maastricht.

Dekker, R.J.P., A. de Grip, L. Borghans, A.G.M. Matheeuwsen, M.H. Wieling, and E.J.T.A. Willems (1993), *Methodiek van het informatiesysteem onderwijs-arbeidsmarkt 1993*, ROA-W-1993/3, Maastricht.

Dekker, R.J.P., A. de Grip, and P.J.E. van de Loo (1990), *ROA-Beroepenclassificatie 1990*, ROA-W-1990/9, Maastricht.

Dekker, R.J.P., A. de Grip, and J.A.M. Heijke (1988), *An Explanation of the Occupational Structure of Sectors of Industry,* ROA-W-1988/2E, Maastricht (also published in *Labour* 4-3, 1990, pp. 3-31)

Eijgenraam, C.J.J., and E.M. Verkade (1988), *BETA; Een bedrijfstakkenmodel van de Nederlandse economie*, CPB, Den Haag.

Evans, G.J., and R.M. Lindley (1973), The use of RAS and Related Models in Manpower Forecasting, *Economics of Planning,* 13-12, pp. 53-73.

Freeman, R.B. (1980), An Empirical Analysis of the Fixed Coefficient Manpower Requirements Model 1960-1970, *Journal of Human Resources*, 5-2, pp. 176-199.

Glebbeek, A.C. (1993), *Perspectieven op loopbanen,* thesis, Groningen.

Granger, C.W.J., and Newbold (1986), *Forecasting Economic Time Series*, Academic Press Inc., Orlando.

Ministerie van Onderwijs en Wetenschappen (1992), *Referentieraming 1992,* Zoetermeer.

OSA (1993), *Vraag en aanbod in het primair onderwijs 1989-1994; Rapport over prognoses en beleidsmogelijkheden*, OSA, Den Haag.

Peeters, H.M.M. (1990), *An Explanation of the Occupational and Educational Structure of Employment by Means of Multinomial Logit,* ROA-W-1990/4E, Maastricht.

ROA (1991), *The Labour Market by Education and Occupation to 1994, Prototype*, ROA-R-1991/5E, Maastricht.

ROA (1992a), *De arbeidsmarkt naar opleiding en beroep en beroep tot 1994*, ROA-R-1992/1, Maastricht.

ROA (1992b), *De arbeidsmarkt naar opleiding en beroep en beroep tot 1994. Statistische Bijlage*, ROA-R-1992/1B, Maastricht.

ROA (1993a), *The Labour Market by Education and Occupation to 1998*, ROA-R-1993/10E, Maastricht.

ROA (1993b), *The Labour Market by Education and Occupation to 1998, Statistical Appendix,* ROA-R-1993/10BE, Maastricht.

Van der Velden, R., and L. Borghans (1993), *Competition on the Labour Market. An Analysis of the Position of Types of Training',* ROA-RM-1993/5E, Maastricht.

Van Eijs, P. (1993), *The Manpower Requirements Approach, Background and Methodology*, ROA-RM-1993/3E, Maastricht.

Van Eijs, P., and L. Borghans (1993), *Manpower Forecasting and the Aggregation Problem: A Microeconomic Approach*, ROA-RM-1993/6E, Maastricht.

Van Eijs, P. (1994), Manpower Forecasting in the Western World: the Current State of the Art, ROA-RM-1994/1E, Maastricht.

Wieling, M.H., A. de Grip, and E.J.T.A. Willems (1990), *Een systematische kwalitatieve typering van arbeidsmarktinformatie*, ROA-W-1990/8, Maastricht also published as 'Forecasting Replacement Demand by Occupation and Education'. International Journal of Forecasting 9, 1993, pp. 173-185).

Wilson, R. (1994), Modelling and Forecasting the Structure of Employment in the United Kingdom, in: H. Heijke (ed), *Forecasting the Labour Market by Occupation and Education,* Boston, pp. 9-35.

Willems, E.J.T.A., and A. de Grip (1990), *Vervangingsvraagprognoses naar beroep en opleiding*, ROA-W-1990/7, Maastricht.

Willems, E.J.T.A., and A. de Grip (1990), *Replacement Demand Forecasts by Occupation and Education*, ROA-W-1990/7E, Maastricht.

Appendix I. Characterizations of predictions and outcomes for expansion demand per occupational class, and the probability of a correct characterization

Occupational class	standard deviation	forecast	outcome	probability of correct characterization: symmetrical interval	probability of correct characterization: forecast error vs. interval width	probability of correct characterization: actual interval
01 Physical scientists and related technicians	0.38	average	low	0.07	0.00	0.07
02 Architects, engineers and related technicians	0.12	average	high	0.23	0.94	0.22
04 Aircraft and ships' officers	0.43	average	very low	0.06	0.00	0.06
05 Life scientists and related technicians	0.42	very low	low	0.11	0.00	0.52
06 Medical, dental, veterinary and related workers	0.18	high	high	0.25	0.68	0.25
08 Statisticians, mathematicians, and related technicians	0.20	very high	very high	0.23	0.00	1.00
09 Economists	0.24	average	very high	0.11	0.00	0.11
11 Accountants	0.52	high	very high	0.09	0.00	0.09
12 Lawyers, public notaries and judges	0.45	low	very high	0.10	0.00	0.10
13 Teachers	0.33	average	low	0.08	0.00	0.08
14 Workers in religion	0.55	low	very low	0.08	0.13	0.08
15 Authors, journalists, and related workers	0.24	low	very high	0.19	0.00	0.19
16 Sculptors, painters, photographers and creative artists	0.40	low	very high	0.11	0.00	0.11
17 Composers and performing artists	0.50	low	low	0.09	0.87	0.09
18 Athletes, sportsmen and related workers	0.55	high	very high	0.08	0.00	0.08
19 Professional, technical and related workers n.c.e.	0.15	low	very high	0.31	0.00	0.31
20 Legislative officials and government administrators	0.65	low	high	0.07	0.00	0.07
21 Senior management, excl. public administration	0.12	high	very high	0.37	0.00	0.36
30 Clerical supervisors	0.24	high	very high	0.19	0.00	0.19
31 Government executive officials	0.52	low	high	0.09	0.00	0.09
32 Stenographers, typists and data-typists	0.30	high	very low	0.15	0.00	0.15
33 Book-keepers, cashiers and related workers	0.12	low	high	0.38	0.10	0.35
34 Computing machine operators	0.45	high	very low	0.10	0.00	0.10
35 Transport and communications supervisors	0.39	low	very low	0.12	0.00	0.12
36 Transport conductors	0.70	low	very high	0.07	0.00	0.07
37 Mail distribution clerks	0.22	low	very low	0.21	0.00	0.21
38 Telephone and telegraph operators	0.42	very low	very low	0.11	0.00	0.56
39 Clerical and related workers n.c.e.	0.26	high	low	0.18	0.00	0.18

Appendix I. Characterizations of predictions and outcomes for expansion demand per occupational class, and the probability of a correct characterization (continued)

Occupational class	standard deviation	forecast	outcome	probability of correct characterization symmetrical interval	forecast error vs. interval width	actual interval
40 Managers (wholesale)	0.48	average	high	0.06	0.43	0.06
41 Managers (retail trade)	0.53	average	average	0.05	0.75	0.05
42 Working proprietors (wholesale)	0.50	high	high	0.09	0.76	0.09
43 Shopkeepers, street vendors	0.41	average	very low	0.07	0.00	0.07
45 Sales supervisors and buyers	0.39	average	low	0.07	0.00	0.07
46 Technical salesmen, commercial travellers etc.	0.38	low	very low	0.12	0.00	0.12
47 Insurance, real estate, and securities salesmen etc.	0.17	high	very high	0.26	0.00	0.26
48 Salesmen, shop assistants and related workers	0.18	high	very high	0.25	0.00	0.24
50 Managers (catering and lodging services)	0.47	very high	very high	0.10	0.00	0.51
51 Working proprietors (catering and lodging services)	0.48	average	very low	0.06	0.00	0.06
52 Housekeeping and related services supervisors	0.49	very low	very low	0.09	0.00	0.57
53 Cooks, waiters, bartenders and related workers	0.34	high	very high	0.14	0.00	0.14
54 Maids and related housekeeping workers n.c.e.	0.33	low	high	0.14	0.00	0.14
55 Building caretakers, charworkers, cleaners etc.	0.32	very high	low	0.15	0.00	0.68
56 Launderers, drycleaners and pressers	0.54	very low	very low	0.09	0.00	0.53
57 Hairdressers, barbers and beauticians	0.47	very low	average	0.10	0.00	0.50
58 Firemen, policemen and warders	0.40	low	low	0.12	0.53	0.12
59 Service workers n.c.e.	0.24	high	low	0.19	0.00	0.19
60 Farm managers and supervisors	0.57	very low	very low	0.08	0.00	0.51
61 Farmers	0.39	low	very low	0.12	0.00	0.12
62 Agricultural and animal husbandry workers	0.36	low	low	0.13	0.72	0.13
64 Fishermen and hunters and related workers	0.68	low	high	0.07	0.00	0.07
70 Production supervisors and general foremen	0.29	very low	very high	0.16	0.00	0.58
72 Metal processors	0.41	low	very low	0.11	0.00	0.11
73 Wood preparation workers and paper makers	0.52	very low	very low	0.09	0.96	0.73
74 Chemical processors and related workers	0.22	average	low	0.12	0.00	0.12

Appendix I. Characterizations of predictions and outcomes for expansion demand per occupational class, and the probability of a correct characterization (continued)

Occupational class	standard deviation	forecast	outcome	probability of correct characterization symmetrical interval	forecast error vs. interval width	actual interval
75 Spinners, weavers, knitters, dyers etc.	0.46	very low	very low	0.10	0.00	0.51
77 Food and beverage processors	0.40	low	very low	0.12	0.58	0.12
79 Tailors, dressmakers, upholsterers, etc.	0.16	average	very low	0.17	0.00	0.17
80 Shoemakers and leather goods makers	0.49	low	very low	0.10	0.00	0.10
81 Cabinetmakers, woodworkers etc.	0.42	low	high	0.11	0.00	0.11
83 Blacksmiths, toolmakers, and machine tool operators	0.19	very low	average	0.24	0.00	0.89
84 Machinery fitters and precision instrument makers	0.12	very low	very low	0.36	0.96	0.74
85 Electrical fitters and related electronics workers	0.28	high	very low	0.16	0.00	0.16
87 Plumbers, welders, sheet metal workers etc.	0.29	low	very low	0.16	0.00	0.16
89 Glass and ceramic workers	0.48	very low	very low	0.10	0.00	0.52
90 Rubber and plastics product makers	0.46	very low	low	0.10	0.67	0.51
91 Paper and paperboard product makers	0.57	average	very low	0.05	0.00	0.05
92 Printers and related workers	0.38	average	average	0.07	0.64	0.07
93 Painters	0.35	very low	low	0.13	0.82	0.50
94 Production and related workers n.c.e.	0.42	very low	high	0.11	0.00	0.55
95 Bricklayers, carpenters, other construction workers	0.16	low	very low	0.29	0.32	0.29
96 Stationary engine and related equipment operators	0.49	very low	very low	0.09	0.60	0.84
97 Despatch workers, packers etc.	0.28	low	very low	0.16	0.00	0.16
98 Transport equipment operators	0.31	low	high	0.15	0.00	0.15
99 Labourers n.c.e.	0.34	low	very low	0.14	0.00	0.14
Total				0.19	0.18	0.25

n.c.e = not classified elsewhere

See notes 19 and 20 regarding minor changes to the occupational class names which were employed in 1989.

Appendix II. Characterizations of predictions and outcomes for replacement demand per occupational class, and the probability of a good characterization

Occupational class	standard deviation	forecast	outcome	probability of correct characterization: symmetrical interval	forecast error vs. interval width	actual interval
01 Physical scientists and related technicians	0.10	low	high	0.41	0.17	0.39
02 Architects, engineers and related technicians	0.07	low	very high	0.55	0.00	0.53
04 Aircraft and ships' officers	0.37	low	very high	0.12	0.00	0.11
05 Life scientists and related technicians	0.11	very low	high	0.37	0.00	0.55
06 Medical, dental, veterinary and related workers	0.06	very low	very low	0.64	0.89	0.50
08 Statisticians, mathematicians, and related technicians	0.09	very low	very low	0.45	0.47	0.89
09 Economists	0.08	low	high	0.52	0.00	0.50
11 Accountants	0.37	low	low	0.12	0.87	0.12
12 Lawyers, public notaries and judges	0.07	very low	low	0.53	0.00	0.66
13 Teachers	0.08	low	very low	0.52	0.33	0.47
14 Workers in religion	0.19	low	high	0.23	0.00	0.23
15 Authors, journalists, and related workers	0.10	very low	low	0.42	0.00	0.74
16 Sculptors, painters photographers and creative artists	0.13	average	low	0.10	0.00	0.10
17 Composers and performing artists	0.16	very low	very high	0.26	0.00	0.54
18 Athletes, sportsmen and related workers	0.16	low	very high	0.26	0.00	0.26
19 Professional, technical and related workers n.c.e.	0.05	very low	low	0.71	0.79	0.62
20 Legislative officials and government administrators	0.17	very high	very high	0.22	0.00	0.61
21 Senior management, excl. public administration	0.06	low	average	0.62	0.03	0.58
30 Clerical supervisors	0.22	high	very high	0.16	0.00	0.16
31 Government executive officials	0.19	high	high	0.19	0.84	0.19
32 Stenographers, typists and data-typists	0.08	high	high	0.42	0.59	0.42
33 Book-keepers, cashiers and related workers	0.04	low	low	0.80	0.63	0.63
34 Computing machine operators	0.31	high	very high	0.12	0.00	0.12
35 Transport and communications supervisors	0.31	low	very high	0.14	0.00	0.14
36 Transport conductors	0.89	very high	very high	0.04	0.00	0.53
37 Mail distribution clerks	0.16	low	high	0.27	0.00	0.26
38 Telephone and telegraph operators	0.30	high	very high	0.12	0.00	0.12
39 Clerical and related workers n.c.e.	0.05	low	low	0.75	0.72	0.61

Appendix II. Characterizations of predictions and outcomes for replacement demand per occupational class, and the probability of a good characterization (continued)

Occupational class	standard deviation	forecast	outcome	probability of correct characterization symmetrical interval	forecast error vs. interval width	actual interval
40 Managers (wholesale)	0.22	very low	low	0.19	0.00	0.56
41 Managers (retail trade)	0.15	very low	very high	0.29	0.00	0.50
42 Working proprietors (wholesale)	0.12	low	high	0.36	0.01	0.36
43 Shopkeepers and street vendors	0.46	high	very high	0.08	0.00	0.08
45 Sales supervisors and buyers	0.08	high	high	0.46	0.82	0.43
46 Technical salesmen, commercial travellers etc.	0.12	low	high	0.34	0.00	0.34
47 Insurance, real estate, and securities salesmen etc.	0.12	low	low	0.36	0.82	0.34
48 Salesmen, shop assistants and related workers	0.12	high	average	0.29	0.46	0.29
50 Managers (catering and lodging services)	0.45	very high	high	0.08	0.00	0.59
51 Working proprietors (catering and lodging services)	0.36	high	very high	0.10	0.00	0.10
52 Housekeeping and related services supervisors	0.34	low	very high	0.13	0.00	0.13
53 Cooks, waiters, bartenders and related workers	0.07	low	low	0.55	0.60	0.52
54 Maids and related housekeeping workers n.c.e.	0.05	low	low	0.72	0.22	0.59
55 Building caretakers, charworkers, cleaners etc.	0.09	low	very low	0.44	0.47	0.43
56 Launderers, drycleaners and pressers	0.27	very high	very high	0.14	0.00	0.57
57 Hairdressers, barbers and beauticians	0.09	low	very high	0.45	0.00	0.41
58 Firemen, policemen and warders	0.10	low	low	0.41	0.63	0.38
59 Service workers n.c.e.	0.07	average	high	0.18	0.38	0.17
60 Farm managers and supervisors	0.29	low	very high	0.15	0.00	0.15
61 Farmers	0.12	low	high	0.35	0.00	0.32
62 Agricultural and animal husbandry workers	0.05	high	average	0.62	0.16	0.54
63 Forestry workers	0.34	high	very high	0.11	0.00	0.11
64 Fishermen and hunters and related workers	0.14	high	very high	0.26	0.06	0.25
70 Production supervisors and general foremen	0.40	high	very high	0.09	0.00	0.09
71 Miners, quarrymen, well drillers and related workers	-	very high	very high	0.00	0.94	0.50
72 Metal processors	0.30	average	very high	0.04	0.00	0.04
73 Wood preparation workers and paper makers	0.21	very high	very high	0.18	0.97	0.77
74 Chemical processors and related workers	0.20	low	average	0.21	0.94	0.20

Appendix II. Characterizations of predictions and outcomes for replacement demand per occupational class, and the probability of a good characterization (continued)

Occupational class	standard deviation	forecast	outcome	probability of correct characterization symmetrical interval	forecast error vs. interval width	actual interval
75 Spinners, weavers, knitters, dyers etc.	0.21	very high	very high	0.17	0.00	0.93
77 Food and beverage processors	0.07	high	high	0.50	0.90	0.46
78 Tobacco preparers and tobacco-product makers	-	very low	very high	0.00	0.00	0.50
79 Tailors, dressmakers, upholsterers, etc.	0.11	very high	high	0.33	0.04	0.56
80 Shoemakers and leather workers	0.30	high	very high	0.12	0.00	0.12
81 Cabinetmakers, woodworkers etc.	0.36	very high	high	0.10	0.00	0.70
83 Blacksmiths, toolmakers, and machine tool operators	0.27	very high	low	0.14	0.00	0.68
84 Machinery fitters and precision instrument makers	0.07	high	low	0.47	0.18	0.46
85 Electrical fitters and related electronics workers	0.15	low	high	0.29	0.21	0.27
86 Broadcasting station and sound equipment ops. etc.	-	very low	very high	0.00	0.00	0.50
87 Plumbers, welders, sheet metal workers etc.	0.13	very high	low	0.28	0.00	0.83
88 Jewellery and precious metal workers	-	low	very high	0.00	0.00	0.00
89 Glass and ceramic workers	0.52	very high	very high	0.07	0.00	0.51
90 Rubber and plastics product makers	0.21	high	very high	0.18	0.00	0.18
91 Paper and paperboard product makers	0.49	low	very high	0.09	0.00	0.09
92 Printers and related workers	0.10	average	average	0.13	0.53	0.13
93 Painters	0.23	high	low	0.16	0.00	0.16
94 Production and related workers n.c.e.	0.27	very high	high	0.14	0.00	0.72
95 Bricklayers, carpenters, other construction workers	0.23	very high	low	0.16	0.00	0.73
96 Stationary engine and related equipment operators	0.28	very high	very high	0.13	0.00	0.52
97 Despatch workers, packers etc.	0.08	low	high	0.53	0.67	0.43
98 Transport equipment operators	0.08	high	very low	0.42	0.00	0.42
99 Labourers n.c.e.	0.16	very high	very high	0.23	0.00	0.93
Total				0.46	0.34	0.47

Appendix III. Forecasts and characterizations of job openings per occupational class

Occupation	forecast	outcome '92	difference	loss	forecast characterization	outcomecharacterization
01 Physical scientists and related technicians	9,000	17,000	−7,000	0.0022	average	average
02 Architects, engineers and related technicians	48,000	133,000	−85,000	0.0274	average	high
04 Aircraft and ships' officers	5,000	5,000	0	0.1044	average	low
05 Life scientists and related technicians	2,000	16,000	−14,000	0.0524	very low	average
06 Medical, dental, veterinary and related workers	83,000	157,000	−74,000	0.0003	average	average
08 Statisticians, mathematicians, rel. technicians	46,000	76,000	−30,000	0.0119	very high	very high
09 Economists	4,000	15,000	−11,000	0.0850	average	very high
11 Accountants	4,000	19,000	−15,000	0.1520	high	very high
12 Lawyers, public notaries and judges	3,000	18,000	−15,000	0.1116	low	very high
13 Teachers	61,000	87,000	−27,000	0.0141	low	low
14 Workers in religion	2,000	5,000	−3,000	0.0002	low	average
15 Authors, journalists, and related workers	2,000	23,000	−22,000	0.2134	very low	very high
16 Sculptors, painters, photographers etc.	6,000	28,000	−22,000	0.0766	low	very high
17 Composers and performing artists	2,000	11,000	−9,000	0.0678	very low	high
18 Athletes, sportsmen and related workers	4,000	10,000	−6,000	0.0545	high	very high
19 Professional, technical and rel. workers n.c.e.	15,000	77,000	−62,000	0.0506	very low	high
20 Legislative officials and govt. administrators	1,000	5,000	−3,000	0.0767	average	very high
21 Senior management, excl. public administration	49,000	157,000	−108,000	0.0577	average	very high
30 Clerical supervisors	8,000	29,000	−21,000	0.2331	very high	very high
31 Government executive officials	5,000	12,000	−7,000	0.0139	average	high
32 Stenographers, typists and data-typists	69,000	46,000	23,000	0.2249	very high	low
33 Book-keepers, cashiers and related workers	70,000	166,000	−96,000	0.0023	average	average
34 Computing machine operators	5,000	4,000	2,000	0.4840	very high	low
35 Transport and communications supervisors	6,000	8,000	−2,000	0.0659	low	low
36 Transport conductors	1,000	3,000	−2,000	0.1876	high	very high
37 Mail distribution clerks	7,000	19,000	−12,000	0.0012	low	low
38 Telephone and telegraph operators	2,000	3,000	−1,000	0.0949	very low	very low
39 Clerical and related workers n.c.e.	136,000	160,000	−23,000	0.0340	high	low
40 Managers (wholesale)	7,000	19,000	−12,000	0.0165	low	average
41 Managers (retail trade)	4,000	14,000	−11,000	0.0963	low	very high
42 Working proprietors (wholesale)	7,000	15,000	−8,000	0.0034	high	high

Appendix III. Forecasts and characterizations of job openings per occupational class (continued)

Occupation	forecast	outcome '92	difference	loss	forecast characterization	outcomecharacterization
43 Shopkeepers and street vendors	29,000	20,000	9,000	0.7915	high	very low
45 Sales supervisors and buyers	16,000	25,000	−9,000	0.0086	high	average
46 Technical salesmen, commercial travellers etc.	11,000	18,000	−8,000	0.0107	low	low
47 Insurance, real estate, and securities salesmen	13,000	29,000	−16,000	0.0114	high	very high
48 Salesmen, shop assistants and related workers	107,000	271,000	−165,000	0.0427	very high	very high
50 Managers (catering and lodging services)	8,000	19,000	−11,000	0.0534	very high	very high
51 Working proprietors (catering and lodging serv.)	11,000	7,000	4,000	0.4541	high	very low
52 Housekeeping and related services supervisors	1,000	7,000	−6,000	0.1758	very low	high
53 Cooks, waiters, bartenders and related workers	37,000	95,000	−59,000	0.0247	high	very high
54 Maids and related housekeeping workers n.c.e.	43,000	103,000	−60,000	0.0022	average	average
55 Building caretakers, charworkers, cleaners etc.	70,000	47,000	23,000	0.1483	very high	low
56 Launderers, drycleaners and pressers	2,000	5,000	−2,000	0.0108	low	average
57 Hairdressers, barbers and beauticians	5,000	22,000	−18,000	0.0584	low	high
58 Firemen, policemen and warders	11,000	24,000	−12,000	0,0001	low	low
59 Service workers n.c.e.	11,000	17,000	−5,000	0.0075	high	average
60 Farm managers and supervisors	1,000	5,000	−4,000	0.1718	very low	high
61 Farmers	27,000	24,000	3,000	0.1237	low	very low
62 Agricultural and animal husbandry workers	31,000	64,000	−33,000	0.0014	low	average
63 Forestry workers	2,000	1,000	0	0.4755	high	low
64 Fishermen and hunters and related workers	1,000	3,000	−2,000	0.0351	low	high
70 Production supervisors and general foremen	14,000	193,000	−179,000	0.3630	very low	very high
71 Miners, quarrymen, well drillers and rel. workers	0	−1,000	1,000	0,0000	average	very low
72 Metal processors	4,000	7,000	−3,000	0.0015	average	average
73 Wood preparation workers and paper makers	1,000	3,000	−2,000	0,0000	very low	low
74 Chemical processors and related workers	8,000	11,000	−3,000	0.0153	average	low
75 Spinners, weavers, knitters, dyers etc.	5,000	4,000	1,000	0.3980	very high	low
77 Food and beverage processors	14,000	27,000	−13,000	0.0026	low	low
78 Tobacco preparers and tobacco product makers	0	−1,000	1,000	0,0000	very low	very low
79 Tailors, dressmakers, upholsterers, etc.	12,000	13,000	−2,000	0.0608	high	low

Appendix III. Forecasts and characterizations of job openings per occupational class (continued)

Occupation	forecast	outcome '92	difference	loss	forecast characterization	outcomecharacterization
80 Shoemakers and leather goods makers	2,000	4,000	−2,000	0.0013	low	average
81 Cabinetmakers, woodworkers etc.	7,000	13,000	−6,000	0,0003	high	high
83 Blacksmiths, toolmakers, machine tool operators	5,000	21,000	−16,000	0.0042	very low	average
84 Machinery fitters, precision instrument makers	20,000	50,000	−30,000	0.0044	very low	low
85 Electrical fitters and related electronics workers	45,000	48,000	−3,000	0.0607	high	low
86 Broadcasting station and sound equipment ops.	0	0	0	0.0000	very low	very low
87 Plumbers, welders, sheet metal workers etc.	42,000	34,000	8,000	0.1496	high	low
88 Jewellery and precious metal workers	0	0	1,000	0.0000	low	very low
89 Glass and ceramic workers	2,000	4,000	−2,000	0.0135	low	low
90 Rubber and plastic product makers	2,000	9,000	−7,000	0.0401	low	high
91 Paper and paperboard product makers	1,000	3,000	−2,000	0.0044	average	average
92 Printers and related workers	13,000	25,000	−12,000	0.0001	average	average
93 Painters	9,000	14,000	−5,000	0.0195	low	low
94 Production and related workers n.c.e.	4,000	9,000	−5,000	0.0015	average	high
95 Bricklayers, carpenters, other construction work	63,000	51,000	12,000	0.1419	high	low
96 Stationary engine and related equipment ops.	−1,000	1,000	−2,000	0.0377	very low	very low
97 Despatch workers, packers etc.	40,000	62,000	−22,000	0.0140	average	low
98 Transport equipment operators	44,000	79,000	−34,000	0.0021	average	average
99 Labourers n.c.e.	19,000	25,000	−6,000	0.1010	high	average

Appendix IV. Characterizations of predictions and outcomes for replacement demand per type of occupation, and the probability of a good characterization

Type of education	standard deviation	forecast	outcome	probability of correct characterization symmetrical interval	forecast error vs. interval width	actual interval
301 General Secondary Education	0.04	average	average	0.29	0.82	0.29
321 Junior agricultural education	0.06	average	high	0.18	0.00	0.18
331 Junior technical education	0.03	average	average	0.34	0.58	0.34
341 Junior transport, communications and traffic education	0.12	low	high	0.32	0.00	0.32
351 Junior medical and paramedical education	0.41	low	very high	0.09	0.00	0.09
361 Lower commercial and administrative educ. (3 types)	0.06	average	very high	0.19	0.00	0.18
381 School for domestic sciences and tech. (excl 361)	0.06	average	high	0.18	0.00	0.18
391 Business security and surveillance training	0.25	low	very high	0.15	0.00	0.15
401 General Secondary Education, int. and higher levels	0.05	low	high	0.70	0.52	0.58
406 Training for driving instructor, sports coach	0.10	low	very high	0.35	0.00	0.35
421 Senior agricultural education	0.06	high	low	0.48	0.19	0.44
431 Senior school for laboratory science	0.28	average	very high	0.04	0.00	0.04
436 Senior technical training	0.03	low	very low	0.89	0.00	0.68
441 Senior education in transport, communications etc.	0.11	high	high	0.29	0.64	0.29
451 Training of nurses and medical receptionists (MDGO)	0.05	low	low	0.62	0.74	0.62
452 Secondary school for medical lab. science, higher level	0.14	low	very high	0.27	0.00	0.27
453 Training for medical clerks etc.	0.31	high	low	0.10	0.27	0.10
454 Training for ward orderlies etc.	0.08	low	very low	0.43	0.69	0.42
461 Senior retail school and intermediate business educ.	0.03	low	low	0.89	0.76	0.50
466 Inter. business educ., dept. of management studies	0.09	low	average	0.40	0.74	0.35
471 Social work and welfare work (MDGO)	0.13	low	low	0.28	0.55	0.28
481 Intermediate health care educ., MDGO/INTAS	0.06	low	low	0.60	0.36	0.52
483 Secondary hotel and catering school, hairdresser's sch.	0.11	high	average	0.29	0.58	0.28
486 Fashion design etc.	0.54	low	very high	0.07	0.00	0.07
506 Primary, pre-primary, and secondary teachers training	0.03	low	very low	0.86	0.62	0.78
511 Training for interpreters and translators	0.33	low	very high	0.11	0.00	0.11

Appendix IV. Characterizations of predictions and outcomes for replacement demand per type of occupation, and the probability of a good characterization (continued)

Type of education	standard deviation	forecast	outcome	probability of correct characterization: symmetrical interval	forecast error vs. interval width	actual interval
516 Training for pastoral work etc.	0.28	low	very high	0.13	0.00	0.13
521 Agricultural college	0.23	low	very high	0.17	0.00	0.17
531 Laboratory college	0.16	low	low	0.24	0.69	0.24
536 Technical college	0.07	low	low	0.53	0.56	0.53
541 Transport, communication and traffic college	0.16	high	high	0.20	0.92	0.20
551 Nursing college, physiotherapy college	0.07	low	low	0.50	0.91	0.45
552 College for medical laboratory science	0.14	low	very high	0.26	0.00	0.25
554 Dietetics college, etc.	0.44	low	very high	0.09	0.00	0.09
561 Business college, excl. admin. and fiscal studies	0.07	low	low	0.50	0.46	0.48
562 Courses for ergonomists, management science HTS etc	0.43	low	high	0.09	0.00	0.09
566 Business science college, legal and admin. studies	0.20	low	very high	0.19	0.00	0.19
571 College of social studies, library studies	0.07	low	very low	0.48	0.68	0.44
583 Hotel college	0.41	average	very high	0.03	0.00	0.03
586 Art academy, academy of dramatic art	0.13	low	average	0.29	0.15	0.29
606 Teacher training (highest level)	0.13	low	very high	0.29	0.00	0.28
611 Language and literature (university)	0.14	low	high	0.27	0.00	0.27
616 Theology (university)	0.20	low	very high	0.18	0.00	0.18
621 Agricultural and domestic sciences	0.31	low	very high	0.12	0.00	0.12
631 Mathematics and physics	0.14	low	average	0.27	0.18	0.26
636 Technical sciences	0.10	low	very high	0.35	0.00	0.33
651 Medical sciences	0.08	low	average	0.46	0.11	0.42
652 Pharmacy	0.47	very low	very high	0.08	0.00	0.50
661 Economics and business administration (B.A.)	0.15	low	high	0.24	0.00	0.24
662 Econometrics, actuary and management (B.Sc.)	0.37	low	very high	0.10	0.00	0.10
666 Law	0.12	low	low	0.31	0.96	0.30
671 Socio-cultural sciences	0.10	low	low	0.37	0.96	0.35
686 Fine Arts	0.56	low	very high	0.07	0.00	0.07

Appendix V. Indicator of the future labour market prospects (LMI) as forecast and recalculated from the actual outcome, with the characterization of both

Type of education	forecast LMI	recalculated LMI	character. forecast	character. recalculation
301 General Secondary Education	0.5	3.5	good	moderate
321 Junior agricultural education	1.1	2.5	reasonable	moderate
331 Junior technical education	1.0	2.8	good	moderate
341 Junior transport, communications and traffic education	−2.1	2.2	very bad	moderate
351 Junior medical and paramedical education	0.1	-	good	-
361 Junior commercial and administrative education (3 types)	22.5	3.4	very bad	moderate
381 Junior home sciences and industry educ. (excl. comm. & adm.)	0.6	3.0	good	moderate
391 Training in business security		4.9	-	bad-
401 General Secondary Education, int. and higher levels	5.3	4.8	bad	bad
406 Training for driving instructor, sports coach	2.8	-	moderate	-
421 Senior agricultural education	0.9	2.4	good	moderate
431 Senior school for laboratory science	1.1	1.6	reasonable	reasonable
436 Senior technical training	1.7	2.0	reasonable	moderate
441 Senior education in transport, communications etc.	0.6	2.2	good	moderate
451 Training of nurses and medical receptionists (MDGO)	1.2	1.9	reasonable	reasonable
452 Secondary school for medical lab. science, higher level	0.7	1.3	good	reasonable
453 Training for medical clerks etc.	2.8	1.9	moderate	reasonable
454 Training for ward orderlies etc.	1.4	1.6	reasonable	reasonable
461 Senior retail school and intermediate business educ.	0.6	1.7	good	reasonable
466 Inter. business educ., dept. of management studies	0.5	2.7	good	moderate
471 Social work and welfare work (MDGO)	2.3	3.1	moderate	moderate
481 Intermediate health care educ., MDGO/INTAS	1.5	2.4	reasonable	moderate
483 Secondary hotel and catering school, hairdresser's school	0.3	2.0	good	reasonable
486 Fashion design etc.	-	-	very bad	-
506 Primary, pre-primary, and secondary teachers training	30.0	2.2	very bad	moderate
511 Training for interpreters and translators	6.4	-	bad	-
516 Training for pastoral work etc.	0.2	0.5	good	good
521 Agricultural college	2.5	–	moderate	–
531 Laboratory college	0.6	1.4	good	reasonable
536 Technical college	1.4	1.5	reasonable	reasonable

Appendix V. Indicator of the future labour market prospects (LMI) as forecast and recalculated from the actual outcome, with the characterization of both (continued)

Type of education	forecast LMI	recalculated LMI	character. forecast	character. recalculation
541 Transport, communication and traffic college	0.3	0.8	good	good
551 Nursing college, physiotherapy college	3.3	2.1	moderate	moderate
552 College for medical laboratory science	1.1	1.1	reasonable	reasonable
554 Dietetics college, etc.	1.3	-	reasonable	-
561 Business college, excl. admin. and fiscal studies	0.7	1.0	good	reasonable
562 Courses for ergonomists, management science HTS etc	0.9	0.8	good	good
566 Business science college, legal and admin. studies	0.9	3.1	good	moderate
571 College of social studies, library studies	8.3	2.5	very bad	moderate
583 Hotel college	0.8	-	good	-
586 Art academy, academy of dramatic art	−7.8	5.7	very bad	bad
606 Teacher training (highest level)	4.8	1.1	bad	reasonable
611 Language and literature (university)	5.0	3.6	bad	moderate
616 Theology (university)	1.2	1.0	reasonable	reasonable
621 Agricultural and domestic sciences	4.1	–	bad	-
631 Mathematics and physics	1.6	2.5	reasonable	moderate
636 Technical sciences	1.6	2.0	reasonable	reasonable
651 Medical sciences	1.4	2.5	reasonable	moderate
652 Pharmacy	0.7	3.0	good	moderate
661 Economics and business administration (B.A.)	2.0	2.1	reasonable	moderate
662 Econometrics, actuary and management (B.Sc.)	0.5	-	good	-
666 Law	4.8	2.3	bad	moderate
671 Socio-cultural sciences	22.5	3.1	very bad	moderate
686 Fine Arts	7.5	-	bad	-